MAR 20 1929

PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

MARCH, 1929

Journal of Educational Research

CONTENTS

	Page
A Brief Method of Evaluating Test Material — *B. R. Buckingham*	159
Promise and Performance in Arithmetic Drill Service — *Guy M. Wilson*	169
An Analysis of the Nature and Frequency of Supreme Court Cases in School Law for the Calendar Year of 1927 — *Richard B. Thiel*	177
A Comparison of the Time Spent by First-Year Students and Expected by Teachers in a State Normal School — *N. A. Wade*	183
Criteria for Judging Thought Content in Written English — *Zelma Langdon Huxtable*	188
A Simple Technique for Determining Whether Children Know the Meaning of Spelling Words — *Carleton Washburne* and *Mabel Vogel Morphett*	196
Editorials	201
Reviews	203
Research Abstracts and Bibliographies	215
News and Communications	224

THE JOURNAL OF EDUCATIONAL RESEARCH is indexed in the International Index to Periodicals

---

Copyright 1929, by Public School Publishing Co.

Entered as second-class matter at the Post Office at Bloomington, Illinois under the act of March 3, 1879

JOURNAL OF EDUCATIONAL RESEARCH

This journal was established to support educational research workers in all phases of the great problems of teaching. Started in 1920 under the editorship of Dr. B. R. Buckingham and associates among the members of an association of directors of educational research, and accepted as the official organ of the Educational Research Association, its purpose has at no time been confined to a narrow line of specialized research but has been concerned also with professional applications for all workers in education.

EDITORIAL BOARD

E. J. ASHBAUGH
Ohio State University

A. S. BARR, Chairman
University of Wisconsin

EARL HUDELSON
University of Minnesota

ASSOCIATE EDITORS

B. R. BUCKINGHAM, *Harvard University*

W. W. CHARTERS, *Director of Bureau of Educational Research, Ohio State University*

M. E. HAGGERTY, *Dean, College of Education, University of Minnesota*

V. A. C. HENMON, *Professor of Psychology, University of Wisconsin*

MURRAY DALMAN, *Educational Consultant, Perkins, Chatten and Hammond, Architects, Chicago, Illinois*

VIRGIL E. DICKSON, *University of California, Berkeley*

CHARLES FORDYCE, *Director, Bureau of Educational Measurements, University of Nebraska*

WILLIAM S. GRAY, *Dean, School of Education, University of Chicago*

KARL J. HOLZINGER, *School of Education, University of Chicago*

ERNEST HORN, *College of Education, University of Iowa*

ARTHUR W. KALLOM, *First Assistant Director, Department of Educational Investigation and Measurement, Public Schools, Boston*

WILLIAM A. McCALL, *Associate Professor of Education, Teachers College, Columbia University*

GEORGE D. STRAYER, *Teachers College, Columbia University*

W. L. UHL, *Dean, College of Education, University of Washington*

GUY M. WHIPPLE, *Secretary, National Society for the Study of Education*

F. J. KELLY, *President, University of Idaho*

EUGENE NIFENECKER, *Assistant Director, Bureau of Reference, Research and Statistics, Public Schools, New York City*

W. J. OSBURN, *College of Education, Ohio State University*

SHELTON PHELPS, *Director of Instruction, George Peabody College for Teachers*

J. B. SEARS, *Stanford University, California*

J. L. STENQUIST, *Director, Bureau of Educational Research, Public Schools, Baltimore, Maryland*

W. W. THEISEN, *Assistant Superintendent of Schools, Milwaukee, Wisconsin*

C. W. WASHBURNE, *Superintendent of Schools, Winnetka, Illinois*

CLIFFORD WOODY, *Director, Bureau of Tests and Measurements, University of Michigan*

THE JOURNAL OF EDUCATIONAL RESEARCH is published at Bloomington, Illinois, monthly, except July and August by the Public School Publishing Company. The subscription price is $4.00; $4.20 in Canada or Mexico; and $4.50 in other foreign countries.

Subscriptions and all communications about reprints or other business matters should be addressed to the PUBLIC SCHOOL PUBLISHING COMPANY, at Bloomington, Illinois.

Address all communications on editorial matters to A. S. Barr, Department of Education, University of Wisconsin, Madison, Wisconsin.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume XIX MARCH 1929 Number 3

A BRIEF METHOD OF EVALUATING TEST MATERIAL

B. R. BUCKINGHAM
Harvard University

NOTWITHSTANDING the development of many new techniques of test construction the problem of determining the difficulty of test items has remained a stumbling block to the construction of better standardized tests. Anybody can see that we cannot get away from a consideration of the difficulty of test questions for the very practical reason that pupils who can answer hard questions are commonly considered to be more capable than those who can only answer easy questions.

It has, however, been pointed out by more than one observer that the weighting of questions introduces more trouble than it obviates. A weight operates as a reward for those who get the question right and a penalty for those who do not. Our attention is usually centered upon the reward rather than the penalty. It is perfectly clear that a child who answers a hard question deserves a large reward; and justice is no doubt served when the large weight attached to this question is added to his score. When, however, a child misses this same question, we withhold from his score the same amount as was added to the score of the successful child. That is to say, we penalize heavily the child who misses a hard question. Correspondingly, we penalize him lightly for missing an easy question. For example, if a pupil cannot name the first president of the United States, we may be expected to reduce our estimate of his ability in history by three points while we may be called upon to reduce our estimate twelve points because he cannot arrange ten events in the order of their occurrence. This is absurd. From the point of view of penalty the child should lose more for failing on an easy question than for failing on a hard one.

Apparently the best solution of this difficulty is to possess ourselves of a considerable number of test items at each appropriate level of difficulty. Penalties and rewards may then be appropriately handled by simply counting the number of items answered correctly at a given level. The pupil may be passed or failed at that level according to a formula

agreed upon—such as, "Pass a pupil at this level if he answers seven questions correctly."

In order to have a respectable number of questions, words, or problems at each significant level of difficulty in each subject, a vastly greater body of test material will have to be evaluated than we now have at our disposal. Moreover, as things are now, an investigator perhaps tries out two or three times as many items as he ultimately uses in his scale or test. The items which he does not use may be rejected because of intrinsic defects as to form or subject matter, but quite as frequently they are rejected because they fail to fall within a convenient segment on a linear projection for difficulty, or because they have not been submitted to a sufficient number of children. In any event the material is usually lost to the educational world.

There are three very good reasons why these conditions exist. The first has to do with the investigator, the second with his coöperators, and the third with the facilities for publishing his work.

In the first place, the plain fact is that the evaluation of a large body of material is uninteresting, tedious, and expensive. It cannot be done with reference to tests as wholes. It must be done by scoring separately each item, no matter how small, to which a response is made by the pupil. If you are evaluating analogies, it is not sufficient to obtain the scores of children on different *groups* of analogies, but you must obtain the scores on each *individual* analogy. Moreover, if you are dealing with pupils of different grades, as is usually the case, the number and percent of correct answers to each question are not enough. A 40-percent question for the third grade means something quite different from a 40-percent question for the fourth or fifth grade. Consequently, your method of scoring must be such as to permit the computation of "grade intervals"—that is, of the amount by which fourth-grade ability exceeds third-grade ability, the amount by which fifth-grade ability exceeds fourth-grade ability, and so on. These grade intervals are functions of the scores of pupils. It therefore becomes necessary for you to adopt a plan for handling the returns so that you will get both question-scores and pupil-scores.

The time-consuming character of this work is appalling. The writer has been made to feel this personally. For example, in the evaluation of words in spelling, he has found that an industrious and capable clerk could not enter upon the necessary double-entry score sheets more than three classes of average size per day if she totaled and balanced the

pupil- and word-scores. When, therefore, one attempts to give each test to as many as a hundred and fifty classes and tries to use fifteen or twenty tests, each having as many as fifty words, he must live long or impoverish himself.

In the second place, a difficulty in the evaluation of test material arises from the natural reluctance of any but the most enthusiastic school people to coöperate in the enterprise. School people want to know how well their children do, not absolutely, but in comparison with other children who have reacted to the same test. Only to a limited extent can an investigator who is merely evaluating new items furnish to the schools this sort of information.

In the third place, not only is there a difficulty due to the tediousness and expense of doing this sort of work and to the reluctance of school people to engage in it, but there is also the fundamental difficulty of securing the publication of a report after one has surmounted these difficulties. There is not a sufficient demand among readers of educational literature for material of this sort. It can only be offered in very long and elaborate tables—tables which no publisher under present conditions can afford to accept for printing in any considerable volume.

Yet it is certain that tables of this sort ought to be available. They should be regarded as source material or material for reference. In this respect they would resemble such tables as are published by the United States Census Bureau. One uses such tables for the primary facts which they contain. Moreover, one may combine these facts in new ways, or one may combine them with facts derived from other sources. Thus, one may make up derived schedules or tables having certain meanings which were by no means apparent in the source material. This, of course, is following the familiar notion in statistics in virtue of which we distinguish between primary and secondary tables. It is probable that we shall have to look to public institutions or to educational foundations for the publication of primary tables of the sort we are considering. The expense involved will no doubt prevent their publication in books or periodicals which have to make ends meet.

The standardization of test items suggest yet another problem. Since the evaluation of data is so difficult is there not some shorter way of appraising test items? We ordinarily think of a measuring device as primarily for the purpose of determining the ability of pupils; but we might think of it as a device to be used in the appraisal of other testing and teaching materials. The present article is concerned with this pos-

sibility. For the sake of simplicity, however, "appraisal" is limited to a single quality, namely difficulty. It is entirely possible that material accurately calibrated for interest or social utility or instrumental value might be used in estimating the importance of new material in these respects. At present, however, we are concerned with the question of difficulty.

Suppose we had a series of evaluated reading selections or stories arranged in such a way that the difference in difficulty between successive selections was equal to the increment in reading ability acquired by typical children in each year of the elementary school. In other words, suppose we had a passage appropriate to the typical reading ability of first-grade children, one appropriate to the typical reading ability of second-grade children, etc. Isn't it clear that such a series of passages could be used not only to test children but also to test reading matter? Could we not, for example, use it in selecting stories for a set of readers?

Perhaps this illustration will serve to indicate what is meant by the idea that test data may be used not only to measure abilities but to evaluate subject matter. If the idea is valid, then it is obvious that this newly evaluated subject matter may in turn be used as test data to assist in securing still more test data; and that in this way the total amount of material of known difficulty may go on accumulating like a snowball rolling down hill. This is perhaps an added reason why research workers should be busy about the accumulation of a reasonably large body of well-selected and thoroughly calibrated elements. The present report relates to this entire question. It gives an account of an investigation in which test data were used in finding the difficulty of additional data. In other words, the report presents a short method of evaluating test material and a criticism of the accuracy of the result.

For a period of two or three years the writer made something of a nuisance of himself by asking his friends, about twice a year, to submit spelling tests. He hopes they gained something of value from the experience, but he is sometimes harassed with doubts. Be that as it may, he obtained by this species of intellectual brigandage a body of material which has been personally useful, but which, for the reasons indicated above, has not been put into print.

On one of these spelling tests returns were obtained from 4264 pupils in eleven school systems. These pupils were about equally divided between the fourth, fifth, and sixth grades. In other words, about 1400 pupils in each of these grades wrote the test. It is assumed that the

percent of correct spellings on each word as thus obtained yields, when transmuted into P. E. values, the true difficulty of each word. In other words, it is assumed that the true difficulty of each of these items was obtained from 4264 observations made in three successive grades of eleven school systems. The probable error of a percent derived from 1400 representative children would not exceed one unit or one percent. It would be much less when derived from 1400 in each of three grades.[1]

Although the test in question consisted of 65 words, only the first 50 of them were used in the present investigation. The following list, consisting of every fifth word, is given in order to show the nature of the test: *restrain, high, purchased, promise, clothes, contains, tried, debts, canvas, lining.* In terms of percents correct the difficulty of the entire 50 words averaged 51 for the fourth grade, 66 for the fifth grade, and 76 for the sixth grade.

All of the 50 words are, by our assumption as to the sufficiency of the testing, of known difficulty. But we are now going to use this known difficulty only with respect to the first 25 words, while for the present we shall regard the remaining 25 words as of unknown difficulty. Let us call the first 25 the *standardized* words and the second 25 the *undetermined* words.

The problem now is to find out the extent to which one may utilize the 25 standardized words to determine the difficulty of the 25 undetermined words. For use in working out this problem, all the available fourth-grade record sheets of City H (one of the eleven cities) were taken from the files. Two schools had been tested at City H; and the fourth grade in one of these schools—here called School No. 1—had 137 pupils, while the same grade in School No. 2 had 65 pupils. Returns were therefore used from 202 pupils.

The problem may be thought of in more specific terms. The writer

[1] The reliability of the "percent correct" for each word may be determined by using the formula,

$$P\ E_p = 0.6745\ \sqrt{\frac{pq}{N}}$$

where p is the percent correct for the word in question, q the percent not correct, and N the number of pupils. For example, with 1400 typical children if the word "restrain" is found to have been spelled correctly by 64 percent of the children, the probable error is a trifle less than one unit or one percent. This is usually explained as meaning that the chances are even that similar results if obtained from other groups of 1400 fourth-grade children would not rise above 65 percent or fall below 63. When the difficulty of the word is secured from not only 1400 fourth-grade children, but also from equal numbers of fifth- and sixth-grade children the probable error will be reduced considerably.

tried to place himself in the position of the superintendent of City H who was supposed to be in possession of the difficulty of 25 spelling words which he had used for test purposes, and which he could no longer use because of the familiarity of the children with them. Such a superintendent might be interested to secure, without going outside of his own system, a reasonably accurate determination of the difficulty of some new words which he could subsequently use as test words. It is submitted that in such a form the problem becomes a sufficiently practical one even though the pursuit of it may be by theoretical methods.

In the first place, the percents correct for the standardized words, as they had been obtained for the three grades of eleven cities, were converted into P. E. equivalents, and these in turn were referred to the theoretical zero of spelling ability. Since the fourth, fifth, and sixth grades were involved, this gave for each word three determinations of its difficulty, each equally valid but carried out upon data for a different grade. Accordingly, these three determinations were averaged to obtain the best determination from the entire data at hand.

The second step was to go through the same procedure using the percents correct for the standardized words as these percents had been obtained from certain fourth-grade classes at City H. Since two schools were involved, a separate determination was made for each of them in order to note how much more reliable the determination from two schools was than the determination from one school.

Brief comment will now be offered on the results as they were obtained for the ten words above mentioned. On the data from School No. 1 the difficulty of the word *restrain* turned out to be 5.56.[2] When data from School No. 2 were combined with those from the first school, the difficulty was found to be 5.88. Now, what we are calling the real difficulty of the word *restrain*—i.e., the difficulty as obtained from the 4200 odd cases in eleven school systems—was 5.84. Thus, the difficulty as obtained from the pupils of one school differed from the true difficulty by 0.28, while the difficulty as obtained from the two schools combined differed from the true difficulty by only 0.04. It will clarify our thinking somewhat if we note, in reference to these differences, that the interval between adjacent columns on the Ayres Spelling Scale is about 0.3 of the units here employed. Accordingly, the discrepancy between the difficulty of *restrain* as obtained from the 202 fourth-grade children at City H and its true difficulty is only 0.012 of a column on the Ayres Scale. This

[2] All numerical statements of difficulty are, unless otherwise stated, in P. E. units.

is remarkably close—indeed, much closer than was the case with most of the other words.

On the word *high* the difficulty from School No. 1 was 5.34, and from the two schools combined, 5.02. The true difficulty was 4.75. As in the case of the word *restrain* we note here the corrective influence of a second school. If the difficulty of the word *high* were accepted as obtained from School 1, it would be taken as 0.59 units (or two columns on the Ayres Scale) harder than is actually the case. When the second school is added, the discrepancy is reduced to 0.27, or about one column on the Ayres Scale.

The word *purchased* is a hard word, but the evidence of its difficulty is revealed just as well by the returns from the two schools at City H as it is in what we are calling the true determination. The difficulty from School 1 is 7.12 and from Schools 1 and 2 combined, 7.19. The true difficulty is 7.12. The discrepancy therefore is only 0.07, which is almost as small as in the case of *restrain*.

The three words on which comment has just been made are in a general way typical of the entire 25 standardized words. From the data for both schools at City H the average difference for each word between the obtained determination and the true determination is +0.25, or somewhat less than the amount known on the Ayres Scale as a column. This means that these fourth-grade children at City H did not perform quite as capably as did the children in general in the eleven school systems.

The amount of the difference is not, for our purpose, the significant thing. The only important question is the question of consistency. If the children at City H—or indeed the children of any school system—are good or poor spellers in general rather than with reference to specific words, then the extent to which their performance on the 25 standardized words differs from the typical performance will likewise be the extent to which their performance on the 25 so-called undetermined words differs from that of typical children. If City H represents the *standardized* words as 0.25 harder than they really are, then it also represents the *undetermined* words as 0.25 harder than *they* really are. Under these circumstances the difference between the record of the fourth-grade children of City H and the record of all the cities on the standardized words gives us a result which we may call the *correction term;* and this correction term may be applied to the record at City H on the undetermined words to obtain an approximation to their true difficulty.

The correction term, as has been shown, is +0.25. The question now concerns the application of this correction term to the word-difficulties obtained at City H for the 25 undetermined words. These word-difficulties when thus corrected may then be compared with the difficulties for the same words that have been found from the 4200 children in the eleven school systems; for, you know, we are only pretending, for experimental purposes, that these words are undetermined.

The reader will recall that 10 of the entire 50 words which we are considering were selected for comment by taking every fifth word beginning with the first. Five of these 10 words are thus among those which are called undetermined. These are, *contains, dried, debts, canvas,* and *lining*. From the 137 children in School No. 1, the difficulty of the word *contains* was found to be 5.59. The difficulty of this word from the 202 children in both schools combined was found to be 5.44. These difficulties may be expected to be too large, because the performance of the children at City H has already been found (i.e., on the standardized words) to be somewhat below the general average. If the children did just as well—no better and no worse in comparison with children in general—on the last 25 words as they did on the first 25, we should expect these two determinations (5.59 and 5.44) to be 0.25 too high. Subtracting this amount (in other words, using our correction term) we obtain 5.34 for School No. 1, and 5.19 for Schools No. 1 and No. 2 combined. The true difficulty of *contains* is 5.55. Thus on this word the standardization obtained from no more than 137 pupils in a single grade and in a single school differs from the true standard by 0.22 P. E. units. This is about two thirds of the interval between two successive columns on the Ayres Scale. It will be observed that on the word *contains* the returns from School No. 1 are in somewhat closer agreement with the standard results than are those of both schools combined. This is exceptional.

On the word *dried* the two determinations from the schools of City H (when corrected as before) are 6.16 and 6.31. The true difficulty is 6.64. The discrepancy for School No. 1 is 0.48 which is unusually large. School No. 2 comes to the rescue to a certain extent and reduces the discrepancy to 0.33, or to about one column of the Ayres Scale.

The word *debts* comes out within 0.2 of the true determination, which is very close. The word *canvas,* however, is "off" by 0.60; and this is the largest difference obtained in the case of any word. *Lining* falls within 0.32 of the true difficulty. Nineteen of the 25 differences between

the corrected values for City H and the true values amount to less than one interval between columns on the Ayres Scale. None of the differences amounts to more than two columns.

When we remember that these results were obtained from only 202 children in the fourth grade of a single city, we may well comment on the closeness of correspondence between the measures of difficulty derived in the manner described and those secured from more than fourteen times as many cases, involving perhaps twenty or twenty-five times as much labor. The real point is not whether we may be satisfied with a standardization based on these particular 200 cases but whether the discrepancy thus obtained does not point to the possibility of using accurately standardized data to reduce very greatly the number of cases required to produce results of satisfactory accuracy. In the judgment of the writer the data obtained at City H justify this conclusion.

Of course, if the proportion of standardized elements is increased in relation to the undetermined elements in a given test the reliability with which the undetermined elements may be evaluated increases. In this experiment half the elements were undetermined. Instead of a fifty-word test composed of 25 words of known difficulty and 25 words of unknown difficulty, the superintendent at City H might have used a test of a hundred words of which 75 were of known difficulty. It is reasonable to suppose that under these circumstances his correction term would have been more accurate and his determination of the difficulty of the 25 undetermined words would have been more satisfactory. Or, supposing that the superintendent had only 25 standardized words at his disposal, he might have contented himself with a test consisting of these 25 standardized words and only 15 undetermined words. His standardization of these 15 words would then be more reliable than the result obtained by the arrangement of test material used in this report. Again, greater accuracy might be secured if the number of children involved in the testing was increased. Instead of 202 children, two or three times that many might have been employed if greater precision were thought to be important. Finally, for the purpose of standardizing the undetermined words, results might have been secured from two or more grades instead of from a single grade.

Thus the question of the possibility of using well standardized material to permit the standardization of other material by a short method is one which lends itself to considerable interesting experimentation. In the first place, one may vary the proportion of standardized to unde-

termined elements; in the second place, one may vary the number of cases; and in the third place, one may vary the number of grades to which the test is submitted. From the results described in this report, when a half-and-half test was used upon 202 children of a single grade, it is inferred that this type of experimentation is promising.

PROMISE AND PERFORMANCE IN ARITHMETIC DRILL SERVICE

GUY M. WILSON
Boston University

MUCH research has been done in the last few years on drill in arithmetic. As a result, new standards and specifications for work in this field have been recognized. This recognition has been reflected in the introduction, prospectus, and advertising material relating to textbooks in this field.

During the past year a sizeable committee in a seminar group undertook to check definitely through a complete analysis of both the prospectus and the materials themselves, the extent to which advertising promises were fulfilled. It is proposed to present herewith a sample analysis of this sort prepared by A. Rebecca Parsons, teacher at Revere, Massachusetts. The analysis presented is that of addition. For convenience, this material will be referred to as "Text Drill Service I."

WHAT IS PROMISED IN TEXT DRILL SERVICE I

Text Drill Service I promises to cover all of the addition combinations and to cover these reasonably in proportion to difficulty as determined by studies of the relative difficulty of the number combinations. There are other promises made with reference to process steps but for the present these we shall neglect. The present task, therefore, is to examine how fully Text Drill Service I meets the requirement of covering all of the addition combinations and giving them emphasis in proportion to difficulty.

HOW MANY ADDITION COMBINATIONS?

As we are to examine how fully the addition combinations have been covered, it is necessary to know how many there are. According to the promises in the text in question, there are 100 and this has been the usual statement in connection with the analysis of addition drill. Years ago it was common to say that there were 45 combinations,—each digit being added to each other digit, making a total of 45. Courtis taught us that the reverses were as important as the direct combinations. Teaching $3 + 4$ does not insure that we have taught $4 + 3$. These reverses added are 45 in number. This, however, neglected the zeros. When the zeros are added we get a total of 100.

But does the above go far enough? Does it take care of column addition? Suppose you are adding up a column and have a total number

of 16, and suppose the next number is 5. Is the fact $6 + 5$? It is not. It is $16 + 5$,—and there is no more assurance that $16 + 5$ has been taught by teaching $6 + 5$ than that $4 + 3$ was taught by teaching $3 + 4$.

It becomes necessary, therefore, in determining the number of addition combinations to consider column addition, and at this point it is necessary to decide how far to go in column addition. Osburn suggested, on the basis of the studies made by others, that extremely long columns were unnecessary. He therefore suggested cutting long columns in two and definitely limiting the decade addition required for column work.

If we fix the limit for column work at $39 + 9$, we will meet fully the requirements that have been indicated by the studies of social and business usage of arithmetic. Taking this as a limit, therefore, how many upper decade combinations do we have? Beginning with 10 we obviously have the possibility of $10 + 0, 1, 2, 3, 4, 5, 6, 7, 8, 9$,—a total of 10 combinations; 11 may have the same; 12 the same; and so up to $39 + 9$. Now from 10 to 39 inclusive there are 30 decade numbers and with possibility of 10 additions for each number we get a total of 300 upper decade combinations. Apparently, therefore, the 100 limit is not quite sufficient. There are at least 300 additional facts needed.

Looking over our arithmetic work we find that addition is not only used in addition but also in multiplication. If we take as a limit of multiplication 9×9, are there any additions called for beyond $39 + 9$? There are. When we multiply 54×8 the first multiplication gives us 3 to carry. The next multiplication 8×5 gives us 40, to which we must add the 3. We therefore have use for the decade addition fact $40 + 3$. How many such facts are there above $39 + 9$ to the limit of 9×9? Any one can easily set these out by doing the multiplications and noting the carrying possibilities. There are a total of 80 as follows:

TABLE I

THE 80 ADDITION FACTS BEYOND $39 + 9$ AS NEEDED FOR CARRYING IN MULTIPLICATION TO 9×9

$40 + 1, 2, 3, 4, 5, 6, 7$	$56 + 1, 2, 3, 4, 5, 6, 7$
$42 + 1, 2, 3, 4, 5, 6$	$63 + 1, 2, 3, 4, 5, 6, 7, 8$
$45 + 1, 2, 3, 4, 5, 6, 7, 8$	$64 + 1, 2, 3, 4, 5, 6, 7$
$48 + 1, 2, 3, 4, 5, 6, 7$	$72 + 1, 2, 3, 4, 5, 6, 7, 8$
$49 + 1, 2, 3, 4, 5, 6$	$81 + 1, 2, 3, 4, 5, 6, 7, 8$
$54 + 1, 2, 3, 4, 5, 6, 7, 8$	

If we now add to the 100 primary combinations the 300 decade combinations needed for column additions and the 80 higher decade combinations needed for carrying in multiplication, *we have as the essential*

basic facts for the mastery of addition, a total of 180. It appears, therefore, that those who constructed the particular Drill Service I, if they made a definite attempt to do it scientifically, started with an obvious handicap in not recognizing more than 100 needed facts in addition. It will be possible to show the actual results.

TEXT DRILL SERVICE I ANALYZED

Miss Parsons has analyzed this drill service so as to show separately the combinations called for in addition alone, those called for in multiplication, and those called for in proving subtraction. It would not be unfair to the authors of the service to show only the combinations in addition. The showing will be a little better, however, if all of the addition work called for not only in addition itself, but also in multiplication and in proving subtraction are given. Since this, therefore, is more favorable to the authors of the drill service, it will be done that way. The reason it is more favorable to the authors is that occasionally a hole in the drill service as originally constructed is filled by mere accident of the requirements in multiplication and subtraction since obviously they were not checked. Tables II, III, and IV which follow herewith show

TABLE II

FREQUENCY WITH WHICH THE PRIMARY COMBINATIONS OCCUR IN TEXT DRILL SERVICE I

+ ↓ →	0	1	2	3	4	5	6	7	8	9
0	8	18	20	26	13	14	16	7	6	7
1	30	97	77	57	56	41	44	44	44	40
2	16	46	61	40	39	25	25	36	48	27
3	22	35	37	41	25	23	48	47	11	19
4	16	35	33	25	44	35	43	19	14	17
5	23	17	34	25	26	55	23	17	18	20
6	23	34	32	32	40	24	52	19	20	16
7	19	22	30	32	17	18	26	42	24	23
8	16	53	55	32	23	23	29	19	50	23
9	27	49	20	21	27	24	29	37	25	60

TABLE III

FREQUENCY WITH WHICH THE UPPER DECADE COMBINATIONS TO 39 + 9 OCCUR IN TEXT DRILL SERVICE I

+ ↓ →	0	1	2	3	4	5	6	7	8	9
10	11	37	28	25	25	39	28	30	31	32
11	5	8	11	8	17	10	11	16	8	20
12	6	27	22	20	24	12	9	9	35	6
13	9	6	11	15	12	5	8	21	13	7
14	3	11	8	8	31	9	26	8	8	10
15	5	15	8	13	18	30	9	11	13	9
16	5	14	17	17	41	13	33	13	6	11
17	7	8	4	22	11	8	14	14	7	9
18	4	14	51	15	12	6	14	12	35	8
19	3	13	6	8	8	8	7	9	6	23
20	3	22	19	18	21	25	29	21	22	26
21	5	4	4	7	5	7	8	2	3	13

TABLE III—*Continued*

22	3	4	6	3	2	4	2	5	19	2
23	2	2	2	7	3	3	7	8	4	3
24	2	11	11	8	15	9	12	2	4	5
25	2	1	4	4	7	13	5	5	3	3
26	1	2	4	4	30	1	17	8	7	3
27	2	4	8	9	4	4	5	12	6	5
28		9	24	9	6	8	5	6	7	6
29		10			3	2	5	4	4	19
30		8	6	7	7	9	19	12	7	6
31	1	1	1	2	2	1	3	1	3	2
32		1	8	5	6	2	2	2	8	2
33				1	1	1	2	6	1	3
34	2	2	2		7	6	2	2	3	2
35	1	2	3	2	4	8	1	3		1
35		4	5	12	9	4	3	2	1	
37		1	1	5	1	1	1	2	1	2
38			2			2	2	3		
39				1	1	3		1	1	7

TABLE IV

FREQUENCY WITH WHICH THE DECADE COMBINATIONS BEYOND 39 + 9 FOR CARRYING IN MULTIPLICATION OCCUR IN TEXT DRILL SERVICE I

+ ↑ —>	1	2	3	4	5	6	7	8
40	3	4	6	5	2	2	3	
42		2	5	3	4	2		
45	1	3	1	2		1	2	2
48	2	1	1	4	4	1	1	
49	1	2	2	2	1	1		
54			2	4	1	2	1	1
56	2	1	3	3	1	4		
63		2	4	1		2		1
64		2	1	1	2			
72	1	2		2	3			
81	1				1	2		

TABLE V

FREQUENCY WITH WHICH THE DECADE COMBINATIONS BEYOND 39 + 9, NOT REQUIRED FOR CARRYING IN MULTIPLICATION AND THEREFORE HAVING LITTLE OR NO JUSTIFICATION FOR OCCURRENCE, OCCUR IN TEXT DRILL SERVICE I

+—> ↑	0	1	2	3	4	5	6	7	8	9
40										
41										2
42										
43			1	2					1	
44		1			1				1	
45										1
46									1	1
47				1				2	2	1
48									2	1
49								1		2
50						1				
51										1
52			1	2	1					1
53					1					1
54										
55										

TABLE V—*Continued*

56							1	1
57					1			
58				1	1		1	
59							1	1
60				1				
61			1	1				
62								
63								
64							1	
65	1					1	2	
66						1		
67				1		1	1	1
68			1	1				1
69			1	1		1		1
70								
71				1				
72								1
73			1					1
74			1			1		
75		1						
76	1							
77								
78					1			
79			1	1		1	1	1
80								
81								
82			1					
83			1				1	
84			1				1	1
85					1	1	1	1
86	1				1		1	1
87					1			1
88				1			1	1
89								
90								
91			1	1				
92				1				
93								
94			1					
95		1						
98	1							
97								
98								
99								

TABLE VI

SHOWING SUMMARY OF COMBINATIONS TAUGHT AND COMBINATIONS USED IN TEXT DRILL SERVICE I

FACTS TAUGHT		FACTS USED	
Primary facts	94	Primary facts	100
Decade facts to 39 + 9	8	Decade facts to 39 + 9	276
Facts for carrying in multiplication beyond 39 + 9	0	Facts for carrying in multiplication beyond 39 + 9	62
Total	102	Total	438
			102
Facts not taught used in examples			336
Facts neither taught nor used			42

respectively,—the primary or first decade combinations; the upper decade combinations needed for column addition to 39 + 9; and the decade combinations beyond 39 + 9 needed for carrying in multiplication to 9 × 9.

Table V gives the decade combinations beyond 39 + 9 which are not needed in multiplication. This table is in itself interesting, as indicating the probability that the authors who constructed the drill service had not themselves arrived at a very fundamental mastery of the process of addition and its requirements.

FINDINGS ANALYZED

Tables II, III, IV, and V tell the story. Careful analysis of the addition work occurring throughout the 186 pages of text and referred to as Text Drill Service I, shows that the authors were of good intent and were, in general, on the right track, but that they evidently did not know the requirements of addition drill. They evidently did not analyze their own drill service as it was constructed. In fact, the indications are that all of the promises of the introduction, prospectus, and the advertising was the work of the advertising department and that it has no necessary connection with the drill service itself. The obvious shortages are further shown from the summary in Table VI.

The above tables indicate that of the total of 480 facts, 102 were taught more-or-less systematically. This means that an effort was made to have the pupils learn 102 facts. The other 378 facts of the 480 required were not definitely or consciously taught. That is, no provision was made for teaching them.

On the other hand, in the construction of the drill service, 438 facts were used. This means that 42 of the 480 facts were neither taught nor used, and that of the 480 facts used in the drill service, 336 were not taught.

ORDER OF DIFFICULTY

While considerable other data are available on the Text Drill Service I, only one other detail will be given, viz., a comparison of the frequency of occurrence of the facts of the drill service with Clapp's *Order of Difficulty* which, the introduction claims, has been used as a guide in determining the frequency of combinations. This is shown in Table VII which follows herewith.

It is interesting to note that the first column of facts, the most difficult, total 497 occurrence, while the column of least difficult facts total

TABLE VII

ADDITION COMBINATIONS IN TEXT DRILL SERVICE I COMPARED WITH CLAPP'S DESCENDING ORDER OF DIFFICULTY

Q 1	Frequency	Q 2	Frequency	Q 3	Frequency	Q 4	Frequency
8+5	20	9 + 0	23	8 + 7	17	7 + 1	16
7+9	21	2 + 6	22	0 + 1	4	2 + 9	21
5+8	17	9 + 3	17	7 + 2	24	2 + 5	24
9+7	32	0 + 6	8	1 + 9	30	2 + 8	43
6+8	18	6 + 5	19	0 + 5	8	4 + 4	37
6+9	13	3 + 8	10	8 + 1	26	4 + 1	23
5+7	10	3 + 4	21	6 + 2	18	3 + 1	21
7+8	20	3 + 9	14	0 + 4	5	4 + 0	14
8+7	17	2 + 3	31	3 + 6	42	1 + 8	28
9+8	24	3 + 5	18	0 + 2	4	9 + 9	57
5+9	15	6 + 3	19	2 + 4	38	7 + 7	40
8+9	21	7 + 3	24	3 + 0	17	2 + 0	10
8+6	18	2 + 7	33	4 + 5	27	6 + 1	18
4+7	18	8 + 4	14	0 + 8	4	5 + 4	19
7+5	13	4 + 8	12	6 + 0	19	3 + 3	37
4+9	14	8 + 0	9	8 + 3	15	1 + 1	69
9+5	22	1 + 0	17	8 + 2	38	9 + 2	11
9+4	23	5 + 2	21	6 + 4	30	8 + 8	48
6+7	15	4 + 2	20	1 + 4	28	1 + 3	37
5+6	18	1 + 2	47	9 + 1	42	1 + 6	25
4+6	35	5 + 3	21	5 + 0	16	1 + 7	23
7+6	21	0 + 3	12	6 + 6	49	2 + 1	32
7+4	13	0 + 5	8	3 + 2	18	2 + 2	52
9+8	20	5 + 1	11	4 + 3	18	5 + 5	49
3+7	39	7 + 0	16	1 + 5	23	0 + 0	5
	497		467		560		759

759 occurrence. This distribution of facts may be justified on the basis of teaching requirements, but it hardly squares with the claims of the authors that the combinations in the drill service are measurably adjusted in frequency to the requirements of Clapp's study on *Order of Difficulty*.

PURPOSE OF THIS REPORT

The purpose of this report is not to condemn a particular drill service. It is to call attention to the fact that drill services can be constructed according to definite requirements and should be so constructed; to the further fact that it is, to say the least, unethical to make claims which are not supported by the facts. It doubtless could be added with safety that for the most part writers of text books and constructors of drill services have done it without having in mind any very definite requirements, have not checked their services, and do not know, when through, very much if anything as to what the drill service contains. It may be suggested also that the time has arrived when superintendents of schools and those who select texts have a right to ask for the

definite specifications with reference to what a book contains. In analyzing the 186 pages in which the process of addition is taught, in order to get an idea of this particular addition drill service, Miss Parsons spent a total of about 200 hours. Is it too much to ask of an author that he do this task before publishing his books?

The tables given in this study suggest a plan for checking addition drill services, which is more adequate than has been common in the past. The present study, however, omits one other essential of a complete check, viz., the distribution of the drill over the process steps.

AN ANALYSIS OF THE NATURE AND FREQUENCY OF SUPREME COURT CASES IN SCHOOL LAW FOR THE CALENDAR YEAR OF 1927

RICHARD B. THIEL
Lawrence College

JUDGE LYON of the Supreme Court of Wisconsin in writing the opinion of the court in a well known and much quoted decision,[1] made the following statement: "Our system of public schools necessarily involves the most delicate relations between parents and children on the one hand and the school authorities on the other, and controversies most frequently arise growing out of the enforcement of school discipline. These controversies, relating as they usually do to the control, management and correction of pupils, are apt to have their origin in wounded parental feelings, and are frequently prosecuted with much bitterness. It is cause for congratulation that so few of these controversies appear in the courts."

While the above quotation is true enough in a way it seems to imply that the major number of controversies pertaining to school matters that reach the state courts of final resort arises out of such disciplinary matters. This may or may not be true. In order to determine the exact nature of these controversies a study was made of all Supreme Court cases involving points of public school law reported for the calendar year of 1927. A more exact statement of the problem of this paper is given below.

Statement of Problem: (1) For the calendar year of 1927, what was the approximate number of cases in public school law passed upon by the supreme courts of the forty-eight states in the union? (2) Are there sectional differences in the frequency with which these cases occur? (3) Is there any relationship between the number of such cases and population, or geographical area? (4) What is the precise nature of the cases passed upon? (5) What are the most likely reasons for sectional variation in the frequency with which cases occur?

Collection of Data: This study is based upon the cases in school law reported by the National Reporter System.[2] The number of cases given is fairly complete and the accuracy of the reports is considered highly

[1] State *ex rel* Burpee v. Burton 45 Wis., 150; 30 Am. Rep. 706.
[2] *National Reporter System Sectional Reports.* St. Paul, Minnesota: West Publishing Co.

reliable, which is reflected by the fact that they are accepted and used by lawyers and courts throughout the nation. The cases were tabulated, first as to frequency by states and sections, and second, according to the points of school law involved. The selection of cases was confined to the calendar year of 1927. The number is deemed sufficiently large to indicate present day tendencies.

TABLE I

TOTAL NUMBER OF SUPREME COURT CASES INVOLVING QUESTIONS OF PUBLIC SCHOOL LAW IN THE FORTY-EIGHT STATES OF THE UNION FOR THE CALENDAR YEAR 1927

Alabama	4	Nebraska	1
Arizona	7	Nevada	0
Arkansas	17	New Hampshire	1
California	11	New Jersey	3
Colorado	3	New Mexico	1
Connecticut	0	New York	4
Delaware	1	North Carolina	7
Florida	2	North Dakota	5
Georgia	14	Ohio	3
Idaho	3	Oklahoma	17
Illinois	5	Oregon	4
Indiana	2	Pennsylvania	7
Iowa	8	Rhode Island	0
Kansas	5	South Carolina	4
Kentucky	17	South Dakota	2
Louisiana	7	Tennessee	5
Maine	0	Texas	17
Maryland	0	Utah	0
Massachusetts	0	Vermont	1
Michigan	3	Virginia	1
Minnesota	6	Washington	2
Mississippi	7	West Virginia	3
Missouri	10	Wisconsin	5
Montana	2	Wyoming	4
		Total	231

From Table I it is seen that seven states have reported no supreme court cases in school law: Connecticut, Maine, Maryland, Massachusetts, Nevada, Rhode Island, and Utah. All of these, except two, are in the eastern part of the United States. Seven states report ten or more cases: namely: Arkansas, California, Kentucky, Missouri, Oklahoma, Texas, and Georgia; all of which are comparatively sparsely populated. That the sparsity of population is associated with the number of court cases is also shown by comparing the sections east and west of the Mississippi River. The nineteen states west of the river reporting supreme court cases show a total of 129 cases, or 55.8 percent of the total, whereas these states make up only 27.8 percent of the nation's population.

Table II shows the number of cases by sections according to the National Reporter System. This division, although somewhat arbitrarily made, follows geographical lines. Here it is again evident that the Pacific and Southwestern states show the largest number of court de-

TABLE II

TOTAL NUMBER OF SUPREME COURT CASES INVOLVING QUESTIONS OF PUBLIC SCHOOL LAW REPORTED IN 1927, CLASSIFIED BY SECTIONS, AND SHOWING THE APPROXIMATE AREA AND POPULATION OF EACH

SECTION	NO. OF CASES	POPULATION (in thousands)	AREA (in thousands of sq. miles)
Atlantic (Conn., Del., Me., Md., N. H., N. J., Pa., R. I., Vt.)	13	17,095	126
Northeastern (Ill., Ind., Mass., N. Y., Ohio)	14	29,411	192
Northwestern (Ia., Mich., Minn., Neb., N. D., S. D., Wis.)	30	13,671	481
Southern (Ala., Fla., La., Miss.)	20	6,905	208
Southeastern (Ga., N. Car., S. Car., Va., W. Va.)	29	10,912	209
Southwestern (Ark., Ky., Mo., Tenn., Tex.)	66	14,574	471
Pacific (Ariz., Cal., Colo., Ida., Kas., Mont., N. Mex., Okla., Ore., Utah, Wash., Wy.)	59	12,699	1,297
Total	231	105,267	2,984

TABLE III

POPULATION PER CASE AND AREA IN SQUARE MILES PER CASE FOR CASES IN SCHOOL LAW IN EACH SECTION OF THE COUNTRY

SECTION	NO. OF CASES	POPULATION PER CASE	AREA IN SQUARE MILES PER CASE
Atlantic	13	1,315,000	9,700
Northeastern	14	2,101,000	14,800
Northwestern	30	455,700	16,030
Southern	20	345,250	10,400
Southeastern	29	376,300	7,200
Southwestern	66	220,000	7,140
Pacific	59	215,200	22,000
Average (Arithmetic mean)		455,700	12,900

cisions. In Table III the ratio of the population to the number of cases is indicated for the United States as a whole. There is one supreme court case in school law to about 455,700 inhabitants. The states in the Northwestern section seem to conform to this average while the states of the Atlantic and Northeastern sections yield a ratio more than three times as favorable and the western states show a ratio of one case to about 218,000 inhabitants which is more than double that of the mean for the nation as a whole.

In making comparisons between the number of cases and the area in square miles, the mean for the nation was found to be one case for every 12,900 square miles. The correlation (Pearson's formula), be-

tween the number of cases and the area in square miles compiled by states was found to be low, only .23, with a probable error of ±.09, and therefore not significant.

In order to determine the precise nature of the issues in the 231 cases passed upon by the courts the classification of American Reporter Key System was used. The Law of Public Schools according to this system, is considered under 170 items, or points of law, each one of which is given a number. Whenever this number is indicated it refers to that particular point of law, for instance, Key No. 18 always refers to Apportionment of School Funds. No. 20 to Nature and Status of School Districts as Corporations, and so on. These 170 items are grouped into eight divisions as indicated in Table IV.

In classifying the various decisions in Table IV the grouping was made not merely according to the number of cases but according to the number of points of school law involved. As some cases involved from one to five distinct points of law, the totals will not correspond to those of Tables I to III.

TABLE IV

POINTS OF SCHOOL LAW PASSED UPON IN 231 STATE SUPREME COURT DECISIONS IN THE FORTY-EIGHT STATES OF THE UNION DURING THE CALENDAR YEAR OF 1927

A. Establishment, school lands and funds and regulations in general	16
B. Creation, alteration, existence, and dissolution of districts	90
C. Government, officers, and district meetings	53
D. District property, contracts, and liabilities	60
E. District debt, securities, and taxation	100
F. Claims against district and actions	11
G. Teachers	38
H. Pupil conduct and discipline of schools	30
Total	398

A glance at Table IV reveals that the first three divisions relate to the territorial and administrative organization of districts and their government. This accounts for 159 of the 398 points of school law tabulated, or 40 percent of the total. Most of these cases may be traced to the sparsely settled regions in which the changes in population call for creating or altering districts, and providing for their administration. Arkansas, California, Georgia, Iowa, Louisiana, Montana, Oklahoma, Oregon, Wyoming, and Texas contribute 78, or practically 50 percent of the total of the 159 points, Texas leading the list with 14.

Controversial matters affecting school property, contracts, district debt, claims and taxation, that is, cases affecting the taxpayers' pocket books, compose 171 of the 398 points, or 43 percent. An inspection of the tabulations reveals that these cases are not confined to any section

but occur rather more frequently in the newer sections, as changes in the boundaries of districts require a readjustment of financial matters, and new districts call for new buildings. It is interesting to note that in this group 33 points of law arose out of contractual liabilities, 34 out of irregularities in issuing school bonds and warrants and 54 dealt with disputes arising out of tax levies.

With regard to the difficulties referred to by Judge Lyon in the relations of teachers to pupils and parents, the number of points of difficulty is only 68, or 17 percent. Of the 38 disputes affecting teachers, three dealt with certification, twelve with matters relating to appointment and contract stipulations, fifteen with suspension, or dismissal, including actions for damages, and the remainder with pensions and other matters.

Among the 30 cases affecting pupil conduct and discipline, seven related to the payment of tuition, seven arose out of claims for transportation, two in regard to vaccination, two to compulsory attendance, so that the number that pertained to "those delicate relations between parents and children on the one hand and the school authorities on the other," was almost negligible. In fact there was only one case that grew out of punishment by expulsion. It is indeed remarkable that in the daily attendance of more than 20,000,000 pupils for the entire year of 1927 there were so few serious legal controversies in matters of discipline and school management. What other branch of our public service can point to a record comparable to this?

This study shows with similar decisiveness that the great storm center in educational litigation is the administration of the school's finances. The number of statutes regulating these matters is common knowledge. The directness and clearness of the wording of these laws is known to all school administrators, hence it is evident that the management of the district's finances is loose, though not necessarily dishonest. Much of this is due to our highly decentralized system of school government. It is equally apparent that the greatest prevention lies in the direction of more capable supervision of school finances by educational administrators and financial experts. The successful school superintendent, more than ever, must be conversant with principles of modern business administration and well informed on legal matters pertaining to his professional relationship.

SUMMARY AND CONCLUSION

1. The total number of Supreme Court cases passed upon in 1927, relating to matters of school law was 231, in which 398 distinct points of school law were involved.

2. There is evidence of a greater frequency of cases in school law in the western and southern states, probably due to a great extent to their sparse settlement and shifting population.

3. The ratio of the number of State Supreme Court cases to the population for the nation is 1 to 455,700; in the eastern states it is about 1 to 2,000,000 and in the western states, 1 to 218,000.

4. The ratio of the number of Supreme Court cases to area in square miles is 1 to 12,900. The coefficient of correlation, (Pearson formula), computed by states between these two factors is .23, with P.E. $\pm$ .09, and therefore not significant.

5. Of the total number of 398 points of school law passed upon, 159, or 40 percent pertained to the territorial organization and administration of school districts. These cases were largely confined to the more extensive and sparsely settled states of the South and West.

6. Forty-three percent of the total number of points of law was concerned with matters of fiscal administration, such as (a) Sale and management of school property, (b) Contracts for buildings, equipment and supplies, (c) School bonds and warrants, (d) Matters dealing with the levy and collection of taxes.

7. The controversies affecting teachers and pupils, including conduct and discipline, were comparatively small in number, about 13 percent of the total. Most of these grew out of disputes concerning the appointment and dismissal of teachers, the payment of tuition, or matters pertaining to pupil transportation; so that the number of cases growing out of disciplinary matters was negligible.

8. The large number of cases growing out of the formation and government of school districts and their fiscal administration suggests the need of further research in matters of educational organization and control, and greater efficiency in fiscal administration.

A COMPARISON OF THE TIME SPENT BY FIRST-YEAR STUDENTS AND EXPECTED BY TEACHERS IN A STATE NORMAL SCHOOL

N. A. WADE
Frostburg, Maryland, State Normal School

THIS study was attempted for the purpose of answering the following questions pertaining to the junior or first-year class of the Maryland State Normal School: How does the average weekly time spent in study compare with the average weekly time that teachers expect? Is the time devoted to study for a recitation in each subject greater or less than each teacher expects? On how many days does the average student study and how is it distributed among those days? What is the average weekly time spent in study, recitation, extra-curricular activities, and assembly combined?

The facts pertaining to students were collected by means of a time-record sheet[1] kept for 6 weeks during the winter term of the school year, 1927-1928. The writer visited each of the three class sections and explained the purpose of the time-record sheets and the mimeographed directions. The records of 80 students—68 girls and 12 boys (Section II)—or 87 percent of the class of 92, were used. To save space only the form for the sixth week, including the horizontal column for the totals for the 6 weeks, is reproduced with the directions. The headings varied according to the subjects pursued by each class section.

The accuracy of the data furnished by the students depends, of course, upon the care and honesty exercised in keeping the time-record sheets. To encourage truthfulness and accuracy the students were promised that no instructor would be permitted to see any of the reports until after he had made his term reports and, furthermore, if they preferred, that they need not return their time-record sheets until they had received their grades.

Without the knowledge of the students, each of the faculty members who was teaching the junior classes was given a sheet with a column headed by his subjects and sections, subdivided into three columns—minimum, average, and maximum—and lined to provide for all the assignments for the six weeks for which students were keeping time-

[1] Modified from the form prepared by Gray, W. S. "Standardization of Student-Teaching Courses." *Educational Administration and Supervision.* VII (1921), pp. 121-132.

TIME RECORD SHEET

Name..... Class Section.....

Date	Days	Arith.		Geog.		Hist.		I. A.		Music		P. E.		R. P.		Sci.		Total	
		P	R	P	R	P	R	P	R	P	R	P	R	P	R	P	R	Hrs.	Mins.
Total in Mins.																			
Total for the 6 weeks	Hrs.																		
	Mins.																		

P —Preparation. Includes anything done in preparation for recitation in each subject.
R —Recitation. Includes time for taking tests during class.
I. A.—Industrial Arts.
P. E.—Physical Education.
R. P.—Reading Process.

1. Each time you work on any subject note the time in minutes. Two cautions: (a) Keep your record accurate and neat. (b) Be sure to include every type of activity which is a part of the preparation in each subject, and the time spent in recitation in each subject.

2. Record the date in the left hand column of the table and insert the number of minutes in the appropriate blank or blanks.

3. At the end of each day add together the number of minutes which appear in the record for that date and express the sum in the right hand column of the table in terms of hours and minutes.

4. At the end of each week add together the number of minutes which appear in the record for that date and express the sum in terms of minutes in the horizontal column of totals for that week.

5. At the end of six weeks add together the number of minutes which appear in the various totals in each column. Record the sum in terms of hours and minutes in the appropriate blanks at the foot of the table.

record sheets. On these sheets the instructors were asked to observe the following directions:

"Before making each assignment in each of the following courses will you estimate in minutes the time—minimum, maximum, and average—that it will require the students to prepare it? Students should not know that this is being kept. Students should not be asked to give the time they spend in preparing each assignment."

An examination of Table I shows that the average weekly time devoted to study by Sections I and III is approximately 2 and 5¾ hours less, respectively, than the average according to the estimate of the teachers; that Section II spends about one hour more per week on the

TABLE I

A Comparison of the Average Weekly Time Spent in Study by 80 First-Year Normal School Students, According to Class Sections, with the Average Weekly Time Expected by the Teachers

(Time expressed in hours and minutes)

Section	No. of Cases	Average Weekly Time for Study	
		Students	Teachers
I	19	21.53	23.51
II	31	21.34	20.36
III	30	20.18	26.00
Class average	(total) 80	21.15	23.29

average than the teachers' estimate; and that, according to the weekly average for the class, the students devote to study 2 hours and 14 minutes less than the teachers expect. It should be noted that Section III, the best group according to the Otis Self-Administering Tests of Mental Ability, reports the least time in study of the three sections while the teachers expect the most time.

Since the students of each section have 20 recitations per week, the average time devoted to preparation per recitation is 66, 64, and 61 minutes for Sections I, II, and III, respectively. Thus the average preparation for a recitation is one hour and four minutes.

TABLE II

The Average Number of Minutes Spent by Each Junior Section in Preparation for a Recitation in Each Subject with Each Teacher's Estimate for the Average Time

Sections	Arith.		Geog.		Hist.		Eng.[1]		R. P.		E. C. E.		Music		P. E.		I. A.[2]		Pen.		O. & P.	
	S.	T.	S.	T.	S.	T.	S.	T.	S.	T.	S.	T.	S.	T.	S.	T.	S.	T.	S.	T.	S.	T.
I	67	60	83	90	84	120	95	90					42	60	27	33	58	95	39	15		
II	70	60					104	105			85	72	38	45	23	30	73	85	45	15	52	30
III	72	70	57	100	78	100			75	90	77	90	36	45	19	40						
For all classes	70	63	70	95	81	110	99	97	75	90	81	81	58	75	23	34	65	90	42	15	52	30

[1] Includes required outside reading.
[2] Includes time devoted to preparing projects.

S. A.—Students' Average.
T. A.—Teacher's Average.

R. P.—Reading Process.
E. C. E.—Early Childhood Education.
P. E.—Physical Education.
I. A.—Industrial Arts.
Pen.—Penmanship.
O. and P.—Observation and Participation.

Is the average preparation for a recitation in each subject one hour? Does each teacher expect the same average preparation? Both questions are answered by the facts presented in Table II, which contains

the subjects pursued by each section, the average time that students spend in preparation for a recitation in each subject, and each teacher's estimate of the average time necessary for preparation for a recitation in each subject. Possibly the most important thing that Table II reveals is the wide variation among the different subjects and sections in both the averages of the students and the teachers. This is true among both the full-credit subjects—arithmetic, geography, history, English, reading process, and early childhood education—and the half-credit subjects—music, industrial arts, physical education, penmanship, and observation and participation. The range for the students' average in the full-credit group is 57 minutes in Section III Geography to 104 minutes in Section II English; for the teachers from one hour in Sections I and II Arithmetic to two hours in Section I History. For the half-credit subjects the range of the students' average is from 19 minutes in Section III Physical Education to 73 minutes in Section III Industrial Arts; for the teachers' estimate from 15 minutes in Sections I and II Penmanship to 95 minutes in Section I Industrial Arts. It should be noted that in industrial arts, a half-credit subject, the teacher expects more time than the instructors in some of the full-credit subjects and in the case of arithmetic receives more. A further study of Table II shows that the students in general spend less average time in study for a recitation in each subject than the teachers estimate.

The typical student not only distributes his study time unequally among the different subjects, but also spreads it unequally among the different days of the five-day week. Moreover, he spreads his study time for the school week over the entire 7 days. He has four recitations per day; but a study of the third week shows that the median time for preparation ranges from 2 hours and 3 minutes on Friday to 3 hours and 55 minutes on Thursday; that slightly less than one-half of the students study on Saturday; and that the majority reports study on Sunday, the median being 2 hours and 3 minutes, approximately the same as Friday.

If to 21 hours and 15 minutes, the average weekly time devoted to study, 18 hours and 12 minutes for recitation, 2 hours for extra-curricular activities,[2] and 1 hour and 50 minutes for assembly per week are added, the typical student of the junior class devotes 43 hours and 17 minutes to four types of school activities. This is 43 minutes less than a standard business week of 44 hours.

[2] Determined by a separate study.

SUMMARY

1. The average first-year student of the Maryland State Normal School spends 21 hours and 15 minutes per week in study, while, according to the average estimate of the teachers, he should spend 2 hours and 14 minutes more, or 23 hours and 29 minutes per week.

2. Based upon the average time per week spent in study, slightly more than one hour is the average preparation for a recitation. But, as a matter of fact, the average preparation varies greatly among all subjects, both full-credit and half-credit subjects. Wide variations characterize the average preparation that the teachers expect in the different subjects, regardless of credit allowed, and in subjects carrying the same credit.

3. The typical junior student scatters his study time unequally over a seven-day week, studying least on Saturday and as much on Sunday as on Friday.

4. Most teachers expect more time for the average preparation than they receive.

5. For 4 types of activities—study, recitation, extra-curricular and assembly—the weekly average is 43 minutes less than for a typical 44-hour business week.

RESULTS

1. The daily schedule has been increased one period per day in order to secure a longer work day at school.

2. The faculty is attempting to arrive at an understanding of what should constitute a reasonable work week in terms of time, to agree upon a more uniform time requirement for subjects giving the same credit toward graduation, and to formulate methods of securing consistent and adequate preparation for each recitation in each subject.

3. During Junior Week good work habits for teachers, how to study, and the effective use of time were stressed more than ever before. Furthermore, the emphasis upon these factors has been continued.

4. A plan is being evolved to extend extra-curricular activities to everybody and to prevent over-emphasis upon the part of some.

5. Students have become interested in budgeting their time and in noting the time required for assignments, some students having voluntarily asked for time-record sheets.

CRITERIA FOR JUDGING THOUGHT CONTENT IN WRITTEN ENGLISH[1]

ZELMA LANGDON HUXTABLE
Edison Junior High, Los Angeles

By what criteria may the thought content of written English be judged? This is a question which confronts and puzzles every teacher of English as well as every teacher of any other subject which calls for subjective written opinion.

Without question, the thought content of written English should determine a large share of its relative merit. Yet in many courses of study the teaching of grammar, mechanics of punctuation, and spelling is emphasized almost to the complete neglect of the thoughts to be expressed. The goal, in other words, has been lost sight of in a near-sighted effort to smooth the way toward it.

During a recent survey of the composition abilities of more than 29,000 children in the Junior High Schools of Los Angeles, an effort was made to evaluate statistically the status of grammar, mechanics of composition, spelling, and thought content. In fact, the particular objective of this work was to discover any evidence which might indicate a greater frequency of certain types of errors at particular IQ levels with a view to the construction of a differentiated course of study for slow, average, and fast learning groups. Miss Mabel Hermans, supervisor of English instruction in the Junior High Schools, has already made use of many of the facts disclosed in this survey in her course of study.

PROCEDURE

In order that the reader may understand the procedure which is described in the paragraphs following, the test instructions given to all English teachers in the Junior High Schools of Los Angeles are explained in full. Each pupil was asked to suppose himself writing to a friend who had previously attended the same school with himself, telling first, about some interesting event which had recently taken place at school; second, about something which would especially interest the friend to whom he was writing, and; third, about his own plans for the summer.

[1] The details of procedure in this problem were worked out under the direction of Dr. A. S. Raubenheimer, of the University of Southern California. The writer is also indebted to Miss Mabel Hermans, Supervisor of English instruction in the Los Angeles Junior High Schools, for permission to use composition papers written under her direction as well as for much personal inspiration and constructive criticism.

Although the composition contents were, to this extent, dictated, the topics were quite within the experience of every child and, consequently, could scarcely offer any impediment to individual habitual modes of written expression. Furthermore, as the assignment was given in the form of a test of which the children had knowledge, its product offers excellent material for a study of this kind where quality and native ability are the variables under scrutiny. Before the test compositions were sent to the Psychology Department, each teacher was requested to read the papers written by her pupils, mark errors with colored checks, and attach to each paper the IQ of its writer.

Inasmuch as a minute examination of more than 29,000 compositions was next to impossible, 1200 papers were finally selected indiscriminately from pupils attending schools in all localities of the city, except for grade and IQ. In other words, 200 papers were taken for each of the six half grades, B7, A7, B8, A8, B9 and A9; of these, 200 for each half grade, 50 papers were chosen from the following IQ groups:

	IQ			
GRADE	89 (and below)	90-110	111-129	130 (and above)
B7	50	50	50	50
A7	50	50	50	50
B8	50	50	50	50
A8	50	50	50	50
B9	50	50	50	50
A9	50	50	50	50

These 1200 papers were then read and evaluated for mechanics, grammar, spelling, and thought content by four graduate students in education who were also experienced teachers.

As might be expected, tabulation of the frequency of errors in mechanics, grammar, and spelling correlated positively, and, on the average, consistently, with the four IQ levels. The evaluation of thought content, however, presented quite a different problem; one, in fact, which is solved with doubtful success by any of the best known composition scales.

That some positive correlation existed between the complexity of the thought expressed and the IQ was immediately apparent to all who handled the compositions from the beginning of the study. But, in spite of the undeniable fact, the task of drawing the line between lower and higher levels of thought content was by no means an easy one. Necessarily antecedent to an attempt to judge thought content with an IQ standard of achievement, one must have at one's disposal first-hand

criteria from many examples of performance at each intelligence level. Hence, the following procedure was outlined and followed.

Six hundred of the twelve hundred papers were read with IQ's known, and descriptive criteria were constructed purporting to classify levels of thought expression to each of the four IQ levels. These "criteria" comprehended five levels of thought complexity which were defined as:

I. Inarticulate Thoughts
II. Unrelated Thoughts (On Sense Perception Level)
III. Related Thoughts (On Sense Perception Level)
IV. Reflective Thoughts
V. Creative Thoughts

Upon each general level of thought expression were found, in each instance, slight differences in excellence due to the absence or prevalence of errors in grammar, mechanics, or spelling. These differences are denoted by the types appearing at each level.

As a second step in procedure, the remaining six hundred of the twelve hundred papers were read and classified by the "criteria" without knowledge of the IQ's of the writers in order to find out whether the first "criteria" had been unduly influenced by the fact that the IQ's were known. No changes were made, however, as a result of this step. In other words, the "criteria" were just as valid when used without knowledge of the IQ's as they were when the information was available. This second step was checked by a study of the IQ's after the papers had been classified by the "criteria."

Finally, as a conclusive check, fifty compositions chosen from all IQ levels of the A9 grade were read and classified for thought content by a group of fifty teachers. These papers were so arranged that each one was read and judged by five different teachers. In all, two hundred-fifty judgments (50 times 5) were obtained. The IQ's on these papers were unknown, and the "criteria" included in this report were used after fewer than five minutes had been spent in explanation of their purpose and application.

Since the significance of the "criteria" depends in large part upon their utility to the layman, the remaining paragraphs are devoted to various features of the conclusive phase of the experiment.

OBSERVATIONAL CORRELATION BETWEEN IQ LEVELS AND LEVELS OF THOUGHT EXPRESSION

Before the presentation of the actual correlation between the teachers' judgments based upon the "criteria" and the IQ levels, a few words may

reveal some interesting features of the observational correlation between IQ's and levels of thought content which became the bases of the "criteria." The unquestionable general correlation between IQ levels and complexity of thought expressed has already been commented upon. Exceptions there were, of course. Indeed, if one's IQ were always revealed in his written thought, English teachers might be relieved of much of the responsibility now charged to them. In reality, on the other hand, the imperative need for improvement in written English, especially with respect to thought content, was emphatically noticeable even among the very bright. Frequently pupils with IQ's of 140 and above fell to average or even low levels of accomplishment judged by the complexity of the thought expressed. In view of the marked decrease of mechanical, grammatical, and spelling errors among the very bright children as compared with the low and average groups, one must conclude that thought content in composition has not been the principal concern of English teachers. A rare individual is he who uses the fine shades of grammar and the mechanics of punctuation to the end of greater, deeper, and more clearly defined thoughts. In spite of much overlapping, as indicated by the "criteria," there were no papers marked with IQ's average, bright, or very bright, which fell to the level of "Inarticulate Thoughts" (See Example, p. 192); and, conversely, there were none marked with IQ's 89 and below which reached a level higher than "Related Thoughts on a Perceptual Level" (See Example, p. 193). It is impossible, apparently, for an individual whose IQ is low, 89 or below, to think reflectively. For him the world and its relationships with himself are perceptual. Specifically stated, his response is governed largely by the stimulus and not by the interaction of the stimulus-percept upon a free concept. Almost invariably the papers from low groups contained gross situations with no analysis of details, and almost as consistently these situations were simply enumerated, failing of both immediate and remote coherence. Papers from average groups tended, also, to reveal a perceptual understanding of environment, but an understanding unmistakably evidenced through consistent relatedness of details. The tendency to unify thought became increasingly consistent with the higher IQ levels.

Examples of compositions agreed upon exactly by four out of five teachers to belong definitely to certain specific levels of the "criteria" follow. A correlation of 63 was obtained from the 250 judgments given by the 50 teachers previously mentioned, and the IQ's distributed in intervals of ten. In fact, a few experienced judges discovered that they

could classify the IQ of a junior high-school pupil correctly within its general placement as low, average, or high, about four out of five times after reading a paragraph or two from his social letter. A composition, judged accordingly, as belonging to "Level II, Type I" might be described as containing unrelated thoughts, with numerous errors in grammar, punctuation, or spelling. The IQ of the writer can also be predicted, if he be of junior high-school age or older, as "low average" or "low." If, as is too frequently the case, a bright child, or even an average child, falls to this level of thought expression he should furnish a specific problem for special consideration on the part of the teacher.

The examples which follow illustrating each thought level and the types upon those levels, are reproduced verbatim up to Level IV, Type 3. From this point on through Level V, the examples quoted are hypothetical. In other words, there were no compositions which reached Level IV, Type 3, or any successive level and type. This may be laid in part to the mental immaturity of junior high-school pupils and in part to the failure of teachers to emphasize the importance of thought in written English.

CRITERIA FOR JUDGING LEVELS OF THOUGHT CONTENT IN WRITTEN ENGLISH

Level I. INARTICULATE THOUGHTS

Type 1. WORDS FAILING TO CONVEY THOUGHT BECAUSE OF MISUSAGE OR MISSPELLING.

Example (IQ 78, Penmanship 20 by Gettysburg Scale)

The Taelaly played the Senior Friday the factaly one the game with a big scaire.

During my summer vaiation Im going to Sandiago I'm study. for Radio engenering.

Type 2. GROUPS OF WORDS FAILING TO CONVEY THOUGHT BECAUSE OF FAULTY MECHANICS OR GRAMMAR, INCOHERENT THINKING, OR MISSPELLING

Example (IQ 70, Penmanship 30 by Gettysburg Scale)

Well since you left school every thing have being change. they took off all the basket ball count and made four daimond to play indoor.

Now that we are going to graduation we all have being pretty busy with are work they are not given out so hard work that you can not do. But to same it seem hard work.

Level II. UNRELATED THOUGHTS on Plane of Mere Sensory Perception

Type 1. ENUMERATED SENSORY EXPERIENCES; UNRELATED TO BOTH IMMEDIATE AND REMOTE CONTEXT; CONTAIN MECHANICAL, GRAMMATICAL, OR SPELLING ERRORS.

Example (IQ 88, Penmanship 30 by Gettysburg Scale)

To-day at school we had what they call boys jinx. It was about the best I seen so far. I will see you in about two weeks for our school lets out June 29. Have you still got your car, if you have we can go to the camp we went last year ....

Type 2. ENUMERATED SENSORY EXPERIENCES: UNRELATED TO BOTH IMMEDIATE AND REMOTE CONTEXT; CONTAIN NO MECHANICAL, GRAMMATICAL, OR SPELLING ERRORS WHICH DETRACT FROM THE MEANING.

Example (IQ 89, Penmanship 70 by Gettysburg Scale)

I am busy studying for tests. I am not graduation but I am going to transfer and I am going to summer school to make up some of my poor grades. There are not many graduation compared with your school. I failed in my Science test and not doing very well in English and Spanish.

Level III. RELATED THOUGHTS on Plane of Mere Sensory Perception

Type 1. SENSORY EXPERIENCES INCONSISTENTLY RELATED; CONTAIN MECHANICAL, GRAMMATICAL, AND SPELLING ERRORS.

Example (IQ 90, Penmanship 80 by Gettysburg Scale)

Last week at school we had the A9 Class Play and next week we are going to have the track meet seventh period on the field. Jack Taylor is the winner of the junior Olymics here and he is going to Atlantic City this summer. We had the A9 class picture taken last Thursday and we will get it today. I sure hope it will be good.

Type 2. SENSORY EXPERIENCES INCONSISTENTLY RELATED; CONTAIN NO MECHANICAL, GRAMMATICAL, OR SPELLING ERRORS WHICH DETRACT FROM THE MEANING.

Example (IQ 112, Penmanship 60 by Gettysburg Scale)

The class play was very good it was called "Four Little Maids of Lavender Time." They gave the play in the afternoon and evening. It was a great success. The cast was made up of the A9 dramatic class.

Whitley Park country club has built a new clubhouse it is very nice. It is a Spanish building and very attractive. The first party in the clubhouse was a great success and all of the members turned out.

Type 3. SENSORY EXPERIENCES CONSISTENTLY RELATED; CONTAIN MECHANICAL, GRAMMATICAL, AND SPELLING ERRORS.

Example (IQ 118, Penmanship 60 by Gettysburg Scale)

You should have seen our class play. Really it was unusually good for Junior High School pupils. It was about a young man who was very poor as a boy but because of his hard work became the Mayor of a big city. I guess there are still chances for us.

Type 4. SENSORY EXPERIENCES CONSISTENTLY RELATED; CONTAINS NO MECHANICAL, GRAMMATICAL, OR SPELLING ERRORS WHICH DETRACT FROM THE MEANING.

Example (IQ 121, Penmanship 90 by Gettysburg Scale)

What are you planning to do this summer? I really only know one thing that I am going to do for sure and that is go to Catalina. I have always wanted to spend some of my vacation there. Last summer I was there for two days but they just seemed to pass so quickly. I have planned to stay about a month there this time. My girl friend in Seattle has asked me to stay with her this summer, so if mother will let me, I am going there for a while also.

Level IV. REFLECTIVE THOUGHTS

Type 1. EXPERIENCE DESCRIPTIVELY ANALYZED; CONTAINS ERRORS IN MECHANICS, GRAMMAR, SPELLING, OR THOUGHT PROCESSES.

Example (IQ 132, Penmanship 50 by Gettysburg Scale)

Last Friday night, "Weddy" and Dan and I went to see our A9

class play. We like it so well that I decided to write to you and tell you about it. I guess you left before you knew the purpose of the class plays. I'll tell you the reason for giving them.

Type 2. EXPERIENCE DESCRIPTIVELY ANALYZED BUT FREE FROM ERRORS.

Example (IQ 145, Penmanship 80 by Gettysburg Scale)

Behold your friend, a wrinkled old woman! That's what I was two weeks ago in a play at our auditorium. It was given by the dramatics class. It was in Chinese fashion, and I was a nurse. You wouldn't have known me with my hair up under a black wig. I had to sob and cry, running all over the stage. It was londs of fun.

Type 3. EXPERIENCE CAUSALLY ANALYZED; CONTAINS ERRORS IN MECHANICS, GRAMMAR, SPELLING, OR THOUGHT PROCESSES.

Example (hypothetical)

For the first time I understood why water appears blue. Perhaps I'm wasting your time to tell you, but I imagine you don't know either, so here goes! Well, you know there are many colors in the spectrum, and that each color has a specific wave length. The color having the shortest length is blue. Now when a ray of sun light passes into a body of water, the blue waves being the shortest of all are absorbed by the particles of matter contained in the water, while the longer waves of other colors are long enough to be reflected back.

Type 4. EXPERIENCE CAUSALLY ANALYZED CONTAINING NO ERRORS.

Example (Hypothetical)

Last week a group of students in our Scince Class went over to Mirror Lake in ..... Park to try an experiment. We wanted to find out just how much light was refracted by different depths of water. It is quite valuable to know this if you are trying to catch fish in clear water, or if you want to recover something from the bottom of any body of water by sight dredging or diving.

Level V. CREATIVE THINKING

Type 1. EXPERIENCE CONTAINING ORIGINAL REFLECTIVE ANALYSIS BUT BEARING THE STAMP OF IMMATURITY.

Example (Hypothetical)

Jane, dear, do you ever feel as I do? Sometimes when I look at the millions of lights down in the valley below us at night, I am frightened and wonder why a thing so insignificant as myself is permitted to understand the meaning of that vast array. And yet, man made those lights and the mighty cities which they illumine. What marvelous cooperation the good Lord inspires in some of his creatures.

Type 2. EXPERIENCE CONTAINING ORIGINAL REFLECTIVE ANALYSIS IN THE FORM OF SPECIAL THOUGHT CONVEYING DEVICES

(Such as figures of speech)

Example

Not far from us here, there is a beautiful little canyon which has been famous or rather notorious, because of the gambling den located a few miles back in the mountains. But last night the ..... Dam broke and a mighty flood swept down the canyon uprooting forests and leaving no trace of the habitat of man. Truly, it seemed as if the gods were reeking their vengeance upon that wicked

and filthy den .... When morning came, the waters were gone, and a deadly stillness hung as a pall upon the path, as it were, of the monster.... At last, the little canyon is cleared of the den of vice but only after man is gone. Indeed, this would be a lovely world if we humans did not mar its beauty.

Type 5. EXPERIENCE CONTAINING ORIGINAL REFLECTIVE EXCELLENCE TO THE EXTENT OF A REAL CONTRIBUTION TO LITERATURE.

Example (Hypothetical)

Mother, perhaps I am antisocial, but there is much in this kind of existence which causes me to question its value as training for manhood's tasks. Last night I wrote these lines after I returned from the Grand Ball.

DISAPPOINTMENT

My heart was bent on gladness,
And my inner being sang;
But the gladness turned to sorrow
And a dirge was all that rang
From the temple of my soul.

These "criteria" are by no means set up as final. There are probably simpler ways of stating the meanings comprehended by them, as, indeed, there are doubtless better ways of classifying the levels of thought complexity themselves. In fact, this is merely an original effort in a field relatively unknown. Just as long as teachers must evaluate subjective written English there will be need for criteria of this general nature, and the sooner some one produces a usable guide, the better.

The writer, in this instance, solicits communication with any person who may use these "criteria" in the evaluation of composition work, not in junior high school alone, but in senior high school and college as well.

A SIMPLE TECHNIQUE FOR DETERMINING WHETHER CHILDREN KNOW THE MEANING OF SPELLING WORDS

CARLETON WASHBURNE
Superintendent of Schools
and
MABEL VOGEL MORPHETT
Director of Research Department, Winnetka, Illinois

IN the grade placement of spelling words their spelling difficulty has usually been made the principal criterion. As a secondary criterion, frequency of use as shown by such studies of adult usage as the Thorndike[1] and the Horn[2] Word Lists has been considered. Some use has also been made occasionally of the so far inadequate investigation of children's written and spoken vocabularies. We have not known, however, which spelling words in the list of any grade were really within the vocabulary of the children who were expected to learn to spell them.

Since the words included in any scientifically constructed speller, such as the *Horn-Ashbaugh*[3] or *Individual Speller*[4], have been chosen on the basis of careful studies of probable ultimate utility to the children who are to learn them, it certainly is desirable that the children should know the meaning of these words. It is palpably absurd to train children to spell a word correctly when they do not know its meaning. On the other hand, it is a distinct waste of time to give children vocabulary drills and training on words which are already perfectly familiar to them. It is, therefore, highly desirable to discover which of the words included in any spelling list are within the vocabulary of the children who are to study this list and which words are unknown to a number of the children and consequently require special vocabulary drill.

It is the purpose of this article to describe a technique which has proved to be both reliable and valid for determining which words in any given spelling list are familiar or unfamiliar to children in the grade for which the list is intended.

The simplest possible method of determining which words are within

---

[1] Thorndike, E. L. *The Teacher's Word Book.* New York City, Teachers College, Bureau of Publications, Columbia University, 1921.

[2] Horn, Ernest. *University of Iowa Monograph in Education.* College of Education, University of Iowa, April, 1926.

[3] Ashbaugh, Ernest, and Horn, Ernest. *Lippincott Horn—Ashbaugh Speller.* J. B. Lippincott Co., 1920.

[4] Washburne, Carleton W. *Washburne Individual Speller.* Yonkers-on-Hudson, New York, World Book Company, and "A Spelling Curriculum Based on Research," *Elementary School Journal,* June, 1923.

the vocabulary of the children is to have the children check on a list the words which they find unfamiliar. Is this method reliable and valid?

To answer this question seven Winnetka fifth-grade teachers[5] took two lists of 37 words each from the four fifth-grade lists in the *Individual Speller*. Three of the teachers dictated Column I to all their fifth-grade children, each child having a copy of this column in front of him at the time of the dictation. Each time a word unfamiliar to a child was reached, that child raised his hand. The teacher, therefore, simply counted the number of raised hands and recorded how many children reported any given word as unfamiliar. The teacher then dictated Column II to the children, but this time she had each child place a check mark opposite any word that he considered unfamiliar, so as to avoid any possible influence of suggestion from the other children. The other four fifth-grade teachers carried out exactly the same procedure except that the children placed check marks opposite unfamiliar words of Column I and raised their hands for the unfamiliar words in Column II.

All checked words were tabulated against the words for which children raised their hands. It was found that, with very few exceptions, the children had indicated the same words as unfamiliar regardless of whether the hand-raising technique or checking technique was used. It was also found that, with great uniformity the same words were indicated as unfamiliar by a fair number of children in each of the seven rooms. Certain words were never reported by a single child as unfamiliar. Other words were reported by a few children in every room as unfamiliar; while still others were indicated as unfamiliar by a considerable number of children in each room.

The uniformity of these results indicated that fifth-grade children were consistent by either the hand-raising or checking technique in indicating the words the meanings of which they knew or did not know. Both these techniques, therefore, appeared to be reliable.

The validity of the checking technique was next investigated. A list of 50 words was so chosen from the fifth-grade spelling list as to include words which were almost sure to be unfamiliar to the children and words which were equally sure to be familiar. The fifth-grade teachers prepared a vocabulary test for each of these 50 words, the test being in the following form:

[5] May Herrick, Helen Sparks, Julia Krenwinkel, Ruth McNamara, Eva Cox, Olga Leap, Ethel Theis.

The climate of a country is (the weather) (the height of the building) (the kind of churches) (the kind of roads) (the way the people dress).

To continue means (to go on) (to stop) (to listen) (to be full). News sent from one person to another is (a wish) (a stage) (a master) (a message) (a bank).

The fifty words for which such tests had been prepared were first mimeographed in a simple list. This list was given to all third-, fourth-, and fifth-grade children in Winnetka (452 children). A similar list of somewhat more difficult words was given to all sixth-grade children (167). The children were asked to check each word with which they were unfamiliar. They were not told that their knowledge of these words was going to be tested. The lists were then collected, and the vocabulary tests on the same words were distributed to the children. On the vocabulary tests, the children underlined the correct answer and the papers were again collected.

All pairs of papers—the checked list and the vocabulary test were turned over to the Research Office for comparison and tabulation.

Table I shows the result of the comparison. It will be noted that

TABLE I

PERCENT OF AGREEMENT BETWEEN CHECKED WORD LISTS AND VOCABULARY TESTS

GRADE	PERCENT OF AGREEMENT	PERCENT CHECKED UNFAMILIAR BUT CORRECT IN TEST	PERCENT NOT CHECKED UNFAMILIAR BUT INCORRECT IN TEST
3	73	13	15
4	80	10	9
5	86	7	7
6	86	6	8

the percentage of agreement is reasonably high from the lowest grade and steadily improves. It will also be noted that in nearly half the cases of disagreement between the checked list and the vocabulary test the child has thought he was unfamiliar with the word and yet has been able to recognize the correct meaning in a multiple choice test. This is, of course, to be expected. Recognition requires less thorough familiarity with the word than does recall. The disagreements under this heading, therefore, can scarcely be considered as true inconsistencies. The last column alone, represented by 15 percent of the third grade, but only 9 percent, 7 percent, and 8 percent in the fourth-, fifth-, and sixth-grades, indicates any true lack of consistency between the children's own judgments as to their familiarity with the words and their actual familiarity as shown by tests.

As a further check on the validity of children's judgments in determining the unfamiliarity of a word, a coefficient of correlation was found between the total number of times a word was marked unfamiliar by children in grades three, four, and five, and the number of times it was found to be unknown in the vocabulary test given to these same children. The coefficient of correlation was .92. Had third grade been omitted this correlation would, of course, have been even higher.

It is therefore evident that it is not necessary to give an elaborate vocabulary test to children in order to determine which of their spelling words are familiar to them. The simple technique of having each child check in his speller those words with which he is unfamiliar is surprisingly reliable and valid.

The technique applied to the *Individual Speller* showed that 11.6 percent of the words in the fourth-, fifth-, and sixth-grade lists were unfamiliar to at least 17 percent and usually 25 percent or more of the children in the respective grades. In the fourth-grade spelling list of 938 words, for example, there are 86 words, which are sufficiently easy to spell to belong in fourth-grade but which should receive special vocabulary drill before the children are asked to spell them.

Every speller should undergo careful analysis to discover unfamiliar words so that vocabulary drill may be provided. The technique is simple and any classroom teacher can apply it. Yet it is surprisingly reliable. At the risk of repetition, the technique may be summarized as follows:

1. Take the children into your confidence. Say something like this: "Before teaching you to spell the words in your speller, I want to be sure you know what they mean. But I don't want to waste your time teaching you the meaning of words you already understand. So as I read these words to you, look at them in your spellers. If you don't know what a word means, raise your hand. Don't pay attention to other children to see whether they are raising theirs, but let me know your own honest opinion. There is nothing wrong about not knowing what a word means. But it is harmful not to admit that you don't know, because then you can't get help. So raise your hand whenever you don't know the meaning of a word."

2. Dictate the words of the entire grade's list, in a series of 10-minute dictations. Pronounce each word clearly, then pause a moment to let the children raise their hands. Record the number of hands raised, in your own desk copy of the speller.

3. Compare, if possible, your marked speller with that of other teachers of the same grade. Pool your results. The more classes included the more generally reliable the data, provided the classes are somewhere nearly alike. A class made up largely of children of foreigners, for instance, would not, however, be fairly rated if averaged in with classes of children from English speaking homes, and vice versa.

4. In general, the meaning of any word that is unknown to 25 per cent of the children had better be taught. A good many other children probably won't know the word any too well. Therefore, prepare a list of "unfamiliar" words from the list for the grade you teach, a word being considered unfamiliar if 25 percent or more of the children don't know what it means. This list should be preserved and used from year to year. It is not necessary to repeat the technique each year, especially if the "unfamiliar" list is made from the pooled results of several teachers' work.

5. Teach the children, each year, the meaning of the "unfamiliar" words in their spelling lists, before you test or teach ability to spell these words. Children should not be taught to spell words which, to them, are meaningless.

EDITORIALS

OBJECTIVE SUPERVISION

The Department of Supervisors and Directors of Instruction of the National Education Association have just published a yearbook[1] on "The Scientific Method in Supervision." This yearbook describes somewhat systematically the several objective devices now available for gathering data about classroom instruction: The attention method, time and frequency activity analyses, teacher's difficulties, and diagnostic uses of tests. The contribution of this yearbook seems to be in its descriptions of these objective data-gathering devices for use in classroom supervision.

Not many supervisors realize the complete inadequacy of conventional methods of classroom supervision. In analyzing teaching situations there is little agreement among supervisors as to what constitutes good teaching and, furthermore the method employed in describing what is observed is highly subjective. The yearbook has made a point of this fact and has proposed a number of devices to improve the situation.

Supervisors, as time goes on, will doubtless come more and more to accept the method of this yearbook, particularly the emphasis upon objective supervision. There is, however, a danger or an inadequacy in objective supervision. Supervision is not *scientific* merely because it is *objective*. It is quite possible that objective supervision may not get all of the facts and that it may misuse the facts gotten. Objective data-gathering such as time and frequency activity analyses merely show what teachers and pupils do. They do not indicate the goodness or poorness of the activity observed. This is done by the supervisor. It is quite possible to have facts objectively collected subjectively interpreted and objective supervision that is unscientific. The point is not that we should avoid objective supervision, but that we should use ob-

[1] "Scientific Method in Supervision," *The Second Yearbook of the National Conference of Supervisors and Directors of Instruction*, New York: Bureau of Publications, Teachers College, Columbia University, 1929.

jective data with care and build up the necessary supplementary techniques to make objective supervision more scientific.

The realization that current supervision is highly subjective has brought forth a whole new crop of data-gathering devises. These data-gathering devices are very much like so many new tests and need to be evaluated as to their validity, reliability, and objectivity. Supervisors will need to exercise extreme care in selecting the data-gathering devices to be used by them in classroom supervision. The fact that someone says that a given device is valuable does not necessarily make it so. A device for collecting data about teaching should be used only when it is shown to bear a definite relationship to good teaching, when the performance is sufficiently constant from day to day to make reliable judgments about it possible, and when the facts collected are sufficiently objective to make it possible for two or more independent observers to arrive at substantially the same conclusion.

The second yearbook of the Department of Supervisors and Directors of Instruction has made a distinct contribution to scientific supervision, particularly in describing new and objective data-gathering devices for use in classroom supervision. Two facts, however, stand out, (and these are pointed out in the yearbook itself); first, that it is quite possible to have objectively collected data and subjectively interpreted, and consequently unscientific supervision; and second, that data gathering devices need to be evaluated for validity, reliability and objectivity.

A. S. B.

REVIEWS

Address all communications relative to book reviews to Editor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

NUTT, HUBERT WILBUR. *Current Problems in the Supervision of Instruction.* New York: Johnson Publishing Company, 1928. xiv + 538 pp.

This book is valuable and suggestive in many particulars.

The Codes of Ethics in Supervision for both the supervisor and the teacher should indicate clearly the coöperative, helpful and human bases for supervision.

The objectives of supervisions are comprehensive, clear-cut and valuable. The plan of definitely applying general supervisory objectives to individual subjects is especially helpful. These might be further improved by greater attention to the child and his interests and needs. Possibly the most valuable phase of the author's objectives in supervision is his reference to teacher growth. He states that "such teachers (those receiving the proper kind of supervision) will grow into more and more expert diagnosticians and become more and more skilled in inventing teaching procedures that meet the complex teaching situation adequately."

The author's discussion of the place of the principal in a supervisory program is in line with the most progressive thought regarding the principalship. The discussion of the types of principals, their selection and promotion points the way toward better principalships. He might have stressed more the specific training for the principalship which progressive professional institutions are now providing. Dr. Nutt's outline of suggestive duties for the principal who is not an expert in certain subjects would be improved by an item calling attention to the principal's function as a supervisor of the technique and psychology of instruction.

The improvement of supervisors in service is an important feature of the book.

Another important element is the classification of plans by size of cities. The plan of organization of some large progressive school systems cannot be applied to smaller places. The book thus shows the current tendencies in various types of situations.

The compilation of duties assigned to principals and supervisors in various school systems is suggestive of a lack of clear-cut conception of supervision and of the democratic methods of conducting coöperative supervision. It also shows a lack of conception of the principal as a director of education in his school and a lack of conception of the supervisor as an expert to assist the principal and the teachers in solving their problems.

The study further shows a type of formalism which rarely results in teacher growth as suggested by Dr. Nutt in his objectives.

The book is a valuable contribution to the field of supervision.

ARTHUR S. GIST
State Teachers College,
San Francisco, Calif.

ROGERS, FREDERICK RAND. *The Amateur Spirit in Scholastic Games and Sports.* Albany, New York: C. F. Williams and Son, 1929. 148 pp.

The author of this little book is one of the leaders in the field of health and physical education who see athletics from an educational standpoint. Believing firmly in the ancient doctrine that *one learns to do by doing*, and finding ample support in the writings of both ancient and modern philosophers and of present-day psychologists for his point of view, he has been earnestly seeking a way by which athletic games could give to the participants the maximum educational value. He believes he has found this by returning the game to the players; by giving the players the opportunity of reaching their own decisions during the progress of the game, and placing upon them full responsibility for these decisions. This is accomplished by taking the coach away from the side-lines and making it impossible for him to dominate by giving instructions through substitutes and signals.

In this little volume, Dr. Rogers has presented the "General Regulation" which expresses this principle which was adopted this past year by The New York State Public High School Athletic Association and tells of the progress made in New York State during the current year. He discusses quite completely the prerequisites to amateurism, quoting freely from philosophers, psychologists and educators in support of the principle. He then presents in Part Two, four most interesting chapters under the general title of The Amateur Spirit—Points of View. These chapters deal with Fair Play, Victory, Competition, and Amateurism. Each of these ideas is discussed clearly with respect to athletics and education.

The volume is presented in the hope that a clearer understanding of what New York State is doing in this respect may influence the school people of other states to likewise render their athletics more educative through pupil instead of coach control. The fact that the principle has worked successfully for one year in a large number of contests, in large cities and small, in sectional and state tournaments, together with the fact that the chief objections are frankly met and well-disposed of from an educational standpoint, should make the volume of special interest to school superintendents and high-school principals everywhere. It might even be hoped that those in charge of college athletics throughout the nation might also evaluate athletics in their respective institutions in terms of the principles set forth.

E. J. ASHBAUGH
Ohio State University

HULL, CLARK L. *Aptitude Testing.* Yonkers-on-Hudson, New York: World Book Co., 1928, xiv + 535 pp.

In this timely volume, the technical procedures of test making and of aptitude prediction are made available for those who can read. With the aid of 95 numbered tables and 60 figures in the text and in addition four lengthy but useful tables in the appendix, all the mysteries are patiently revealed, step by step. It is written for the layman, yet the specialist will find much of value in its closely packed pages for its methods have general application. The author takes the empirical view of testing: Anything is a good test which can be reliably and cheaply enough measured in the light of its contribution to a battery for predicting an independent criterion of the aptitude in question. And the theoretical aspects are made clear in the only treatise known to the writer which

does so—group factors or determiners, dice throws or chance, correlations among determiners—all elucidated with a little simple algebra. The several chapters on "anatomical and other alleged signs of aptitude; the basic constitution of aptitudes and tests; combining the tests to secure the maximum forecasting efficiency"—contain permanent contributions to the field of testing.

The reviewer regrets that more references to the original sources of the formulae were not provided; (it contains a bibliography of 105 titles); also that the author refrained from speculating as to the nature of aptitude "determiners" although, obviously they are but means to an end. Human aptitude acts "*as if* it were composed of, or dependent upon, these more or less hypothetical entities." A few typographical errors have crept in, none of which, read in connection with the numerical examples given, will lead to serious error with the possible exception that in formula (15), page 397, the prime sign should be attached to the 0 of the numerator rather than to the denominator. As a rosy, but eminently practical, prophecy of the future, we shall some day make aptitude forecasts automatically on a machine which grinds them out while you go to dinner and then stops automatically when finished! Three of the machines have been made and they work as described.

HERBERT A. TOOPS.
Ohio State University

MONROE, WALTER S., AND WEBER, OSCAR F. *The High School.* New York: Doubleday, Doran & Company, Inc., Teacher Training Series. 1928.

This is the second book of a series of three volumes planned by Dr. W. S. Monroe as a coördinated series of texts in the field of education. This volume has been written primarily for high-school teachers and is designed to give an acquaintance with the significant facts about high school as a unit in our educational system. The scope of the book is suggested in the following sentence with which the authors open their first chapter: "What should the teacher know about the high school as an educational institution?" There are single chapters on "The Development of Secondary Education in the United States," "Secondary Education in England, France and Germany," "The Downward and Upward Extension of the High School," and two chapters on the objectives of the secondary school. The problems of the curriculum are given more extended treatment than other problems, and the materials included are chosen so as to help the teacher "maintain the right attitude toward proposals for changes and be able to evaluate them wisely." The topics of "Extra-curricular Activities" and a few of the more important administrative problems are given brief treatment.

In writing the book the authors have drawn very extensively on special studies and the reports of national committees. The treatment throughout is largely historical. The learning exercises at the close of the chapters constitute one of the most helpful features of the book. These present a sufficient variety of problems and questions to make it comparatively easy for a skillful teacher to encourage wide reading and to stimulate class discussion.

In all probability, there are some who would criticize the book on the score that it presents only a survey of materials from the fields of history of education, philosophy of education, the curriculum, administration and guidance. It is my opinion, however, that the authors should be commended for undertaking the preparation of a text for an intro-

ductory course. We need the highly specialized books for our advanced students, but our beginning students need texts that will give them a professional outlook upon the essential problems of the larger units of our educational system.

J. B. EDMONSON
University of Michigan

HOLLINGWORTH, LETA S. *The Psychology of the Adolescent.* New York: D. Appleton & Co., 1928. 253 pp.

According to the author of this book, it was written to make the intelligent youths who are actually in the midst of the adolescent adjustments described realize that the experiences they are going through are both normal and universal, and to give to parents and teachers who have the responsibility of directing these young people a clearer knowledge of the adjustments through which "the boy or girl in the 'teens' is passing." With this object in mind, and endowed with the belief that the right sort of help can only be given by parents and teachers when adolescence is foreseen in the rearing and education of the infant and child, the author proceeds to discuss for parents, youths, and teachers what she regards as the universal problems of this period of development.

What the author of this book regards as the outstanding problems of the adolescent period of development may be sensed by quoting the chapter headings under which these problems are discussed; namely, "Psychological Weaning," "Seeking Self-Support," "Mating," "Achieving a Point of View," "The Meaning of Maturity," and "Finding the Self." These problems are discussed in a superficial but entertaining way that will prove helpful to parents and many young people. But the serious student of Genetic Psychology will experience great disappointment in this author's treatment of this unique and important period in human and individual development.

No attempt is made to present the facts recently ascertained about this stage of development in the light of what has already been determined about the laws and principles of individual development by previous investigations in the field. That many new methods and techniques have been developed during the past twenty-five years that might be used to study the problems in this field and that much quantitative material has already been collected that bears directly on this important stage of development the author of this text is aware, but little of this material is assembled and interpreted in the light of previous investigations. The book is, in fact, a mere formulation of some of the problems of the adolescent, "as they appear at present, under conditions of contemporary life." As such, it will doubtless be valuable to teachers and parents who seem to be adverse to reading or studying anything that is scientific or thorough.

The author has, therefore, gathered together some data, set forth a few historical facts and brought to mind some of the problems of adolescence in an effort to shock the "guardians of youth" into a realization of the seriousness of these problems. In this she has succeeded, for her style is both clear and entertaining. But as a text in the Psychology of Adolescence, the book should be classed as a failure.

WILLIAM FREDERICK BOOK
Indiana University

SNEDDEN, DAVID. *Educational Sociology for Beginners.* New York: The Macmillan Company, 1928. xiii + 636 pp. We might expect one who has entered

educational sociology from the field of education to give particular attention to educational problems; but Dr. Snedden in his book for beginners has been primarily concerned with social problems. At least three-fourths of his more than 600 pages is taken up with sociological discussions. Not even Dr. Ross has introduced into his writings a wider range of illustrative materials which throw light upon the moving currents of contemporary civilization. In fact, there may be some danger lest the wealth of allusion and shrewd comment may tend to bewilder rather than clarify the thinking of many of the unlearned for whom it was written.

Every chapter in the book is stimulating and illuminating. As in all of Dr. Snedden's recent work, the content is intensely analytical. It is question-raising rather than problem-solving. In the present state of educational sociology this may be the better technique; but it involves considerable repetition and one wonders whether the beginner might not gain more help from a more synthetic and constructive treatment of his actual problems than is here attempted. It is true that many students are being offered courses in educational sociology who have not had previous work in general sociology. Yet the growth of sociological study in normal schools, as well as in colleges and universities, has been so rapid that the present writer feels the time has arrived when the educational sociologist can devote his time as fully to the sociological treatment of educational problems as the educational psychologist does to the psychological treatment of the same or similar problems.

Among the several texts recently published none are more thought provoking than Dr. Snedden's and it will, without question, find a useful place in educational classes. It is well equipped with problem questions and suggestive readings.

WALTER R. SMITH
University of Kansas

MCALLISTER-OTIS. *Child Accounting Practice*. Yonkers-on-Hudson, New York: World Book Company, 1927. xii + 196 pp.

"Child Accounting Practice" discusses in each of its four units a distinct problem in accounting for the pupils of a school system. The first unit raises the general question as to what is the purpose of such work; it discusses general principles that pertain to the field; and explains in detail the purpose and nature of the McAllister-Otis Child Accounting System. Unit Two describes for the teacher her part in child accounting. Sample forms are shown, their purposes are discussed, and the detailed techniques needed in using the forms are described in great detail. Such tasks as keeping the class roll, recording attendance, recording scholarship, recording test scores, making permanent cumulative records, and making reports are considered.

Unit Three describes the principal's part in child accounting. Samples of those forms which are needed by the principal are shown; these are records of character ratings and of health, and reports which are to be sent to the superintendent's office. Such tasks as supervising the work of his teachers in child accounting, record ratings on character traits, recording health data, summarizing the reports of his teachers, and making the "Administrative Report" to the superintendent's office are described in detail. Unit Four outlines the work of the central office in relation to child accounting. The work involved in providing an accurate school census, in

keeping track of individual pupils and in filing school records and reports is analyzed for the reader.

The authors have set a standard of excellence which writers in the field of child accounting cannot ignore in the future; they have not been content to name a series of record and report forms and call such a series a system; they have developed each form in relation to the other forms of the system; they have provided a detailed description of how each unit of their system is to be used by teacher, principal, and central office; and they have shown why they believe each form in the system is necessary. Most systems of child accounting forms seem to assume that the technique of using the forms is entirely obvious from an examination of the forms themselves. Until this misapprehension is displaced and until formulators of child accounting systems refuse to leave a definitely technical procedure to chance, child accounting practices will continue to be non-uniform, inaccurate, and in many ways of little worth. It is with pleasure that I review a book that has made a contribution of this nature.

It seems almost out of place to offer a criticism where such a contribution has been made. If I may be indulged, however, I will offer three—one minor and two major. The authors suggest that "whenever a pupil has been absent more than six consecutive half days, he should be dropped from the roll." For a book with so many progressive features in it, it seems a bit disappointing that this rather vicious rule governing the recording of attendance should be given further encouragement. The authors at least partly justify the rule upon the basis of the old recommendation of the N. E. A. Committee report on Uniform Records and Reports of 1912. The more recent report of the N. E. A. Committee on Uniform Records and Reports of 1928 repudiates this rule. In justice to the authors, the latter report was published after their book came from the press.

"Child Accounting Practice" is undoubtedly intended as a handbook for the teacher. The authors say, "Obviously you (addressing the teacher) will find it of advantage also to have one copy of this Manual. Your careful study of it will make your child accounting work a delight to you . . ." It is questioned whether the teacher with little or no training in child accounting will read easily the somewhat technical descriptions of forms and procedures. New terms, a number of symbols, and a continual reference to additional pages and to footnotes for a proper understanding of given points, at times makes the reading of the manual a little difficult. As a guide for the technician in child accounting the work is excellent.

Again one is slightly disappointed when the items upon the cumulative pupil records are reviewed. Increased emphasis upon guidance from the seventh through the twelfth grades demands a wealth of detail about the individual. The system would seem to stress the counting aspects of child accounting somewhat at the expense of its guidance aspects. In spite of these possible criticisms, the book is a decided contribution to the field of child accounting and the reviewer hopes that McAllister and Otis will continue their contributions to this field.

A. O. Heck

Ohio State University

New York Society for the Experimental Study of Education. *Contributions to Education*, Volume II. Part I, "Experiments in the Classroom," edited by J. Carleton Bell and

and Part II, "Problems in Teacher Training," edited by Ambrose L. Suhrie. Yonkers-on-Hudson, New York: World Book Company, 1928. x + 425 pp.

This volume, like Volume I, is a tribute to the inspirational and directive influence of Professor J. Carleton Bell, founder of the New York Society for the Experimental Study of Education, and its executive secretary during the ten years in which its membership has grown from 74 to 1500. Twenty-six accounts of studies and observations carried on by classroom teachers are reported in Part I. In nearly all of these studies at least one standard intelligence test or educational measure has been employed.

With a few exceptions, the papers present little or no contribution to technique or scientific knowledge; nor do the authors pretend to make such contributions. These studies do, however, represent a real step forward since they were carried on by classroom teachers, and in most cases by teachers who ordinarily are considered laymen in educational research. Some of the most valuable research of the future is, doubtless, destined to emanate from this source, as Buckingham has strongly claimed in his editorials and in his book, *Research for Teachers*. Such studies will probably do more to stimulate an experimental attitude on the part of the teachers than will the more refined researches from the laboratories of skilled technicians. The spread of an epidemic among classroom teachers of searching for objective information regarding how their children learn, their specific difficulties, failures, and successes, will result in marked improvement in the art of teaching.

Part II is a collection of papers read at a conference, descriptive of the various phases of work carried on by the leading eastern teacher-training institutions. From each of several such institutions reports are made of the practice in that school concerning such subjects as methods of selecting students for admission, plans for helping students make up deficiencies discovered in their preparation after admission, content and conduct of courses in the "introduction of teaching," the teacher's personality, student teaching, procedure for eliminating the unfit, appointment and supervision of beginning teachers, student government, and student-faculty coöperation. Dr. Suhrie, who edits Part II, reveals the vision which has always characterized his efforts in teaching and school administration.

GARRY C. MYERS
Cleveland School of Education

FITZPATRICK, EDWARD A., AND HUTSON, PERCIVAL W. *The Scholarship of Teachers in Secondary Schools*. New York: Macmillan Company, 1927. x + 208 pp.

Two essays, quite different in character, which were presented in competition for the Sachs Prize of 1926 are found in this book. Together they constitute probably as comprehensive and as up-to-date a discussion on the scholarship of secondary-school teachers as one can find. The winning essay by Dean Fitzpatrick surveys briefly the essential factors bearing upon the scholarship of teachers and then attempts in a constructive and fundamental way to present a new policy in teacher training. It finds in the German system much that can be adopted profitably in this country. Hutson's contribution, on the other hand, is quite largely a factual study with concisely stated recommendations based

upon facts laboriously accumulated. The difference between the two studies is well illustrated when each considers the functions of teacher-training schools. Fitzpatrick asks and attempts to answer the question, "What should the teachers colleges do?" Hutson examines the catalogues of training institutions in order to determine what they are doing. Both recognize, however, the present patchwork preparation as an evil, and thus both come face to face with the necessity for a more liberal professional training. This prompts Hutson to study the present relationship between the courses taken in college by teachers and the subjects which they teach. It prompts Fitzpatrick to inquire as to what constitutes a liberal preparation for teaching. In this way each writer approaches somewhat differently the problem of professionalizing subject-matter for the training of teachers. Neither, however, bores as extensively as he might into the rich possibilities of this question.

Space will not permit us to compare fully the points of agreement and differences in the two discussions. Both admire the California requirement of one year of graduate work for secondary-school teachers; both see the continuation of the small high school as a constant source of infection; both insist that improvement is absolutely dependent upon a better coördinated and more concentrated type of training prior to teaching in the subjects to be taught; and both insist that teachers be competent to give instruction in three high-school subjects.

V. T. THAYER

Ohio State University

WOOD, BEN D. *New York Experiments with New-Type Modern Language Tests*. Vol. I, Publications of the American and Canadian Committees on Modern Languages. New York: The Macmillan Company, 1927. xxii + 339 pp.

As volume one of the above important series of publications Professor Ben D. Wood's "New York Experiments with the New-Type Modern Language Tests" has particular significance for the heads of schools and colleges as well as for linguists and educationalists. It reveals in a striking manner a splendid collaboration between the administrative officers of the school system in a great metropolitan center and in an influential state and the corps of investigators without which no worth while research on a large scale is possible. This is a praiseworthy example that might profitably be imitated by other cities and state systems. Also by the personal method of inquiry employed and by the frank presentation of its discoveries, it seems to promise that the whole trend of the investigation by the *Study* will be away from the empirical method in the direction of the scientific. This is as it should be, because heretofore altogether too many of our opinions and decisions about matters pertaining to our field have been established on subjective rather than on objective grounds and hence to be deprecated.

As shown by the title page, Professor Wood's monograph really comprises three different experiments, namely, (1) a Survey of Modern Language Achievement in the Junior High Schools of New York; (2) the Regents Experiment of June, 1925, with the New-Type Tests in French, German, Spanish and Physics; and (3) a Second Survey of Modern Language Achievement in the Junior High Schools of New York City, June, 1926. Some of the principal features of each report are summarized under their respective heads as follows:

PART ONE

The experiment in the Junior High Schools of New York in June, 1925, was conducted with what are now known in their revised form as the American Council Beta French and Spanish Tests, about 25,000 copies of which were employed on this occasion. The author devotes the greater part of five chapters to explaining the construction and validation of the tests, setting forth the results obtained with the same, making an analysis of the difficulties of the questions, reproducing the tests themselves, and recommending certain remedial measures in the light of his findings. In regard to the first point the reader is informed that the new tests were made after four years experimentation and consist of a vocabulary part of the multiple choice type, a comprehension part also of the multiple choice type, and a grammar part of the completion type, there being 220 items in the French test and 225 in the Spanish. The reliability coefficient of both tests, each based on 2000 cases, stepped up by the Spearman-Brown Formula is .97 as compared with .80, the estimated maximum reliability for the old-type modern language examination, according to Professor Wood. The author points out further that the error of estimate of the above new type test expressed in terms of standard deviation is about one-third that of the old-type examination with a reliability of .80.

Lack of space forbids the discussion of more than one or two of the most important revelations resulting from the testing. Perhaps the most impressive of these findings is the overlapping of classes, for the tests constitute a language yardstick, so to speak, which allows comparable measurements to be taken. In this connection the author writes thus:

> From the graphs for all the French classes, therefore, we learn that over 2000 of the nearly 19,000 junior high school French students are misplaced by a whole semester or more, and that approximately 10,000 of them are nearer to a higher or lower class-average than to the average of the class in which they were reciting during the spring of 1925. (p. 26.)

After demonstrating that this same heterogeneity of classification exists in the different semesters of Spanish the investigator traces back the origin of the trouble to the schools, which differ at times from one another in standards of work a half year and even a whole year. Results like these show the need of sound classification of students based on achievement in a subject rather than on units of time and the old subjective form of examining and grading. It is to the credit of those in charge of the modern languages in the junior high schools of New York that a reclassification of all students that took these tests was drawn up and a laudable attempt made to put it in force. We are told, however, that the full advantage of the new plan could not be taken because of administrative exigencies and difficulties of promotion and demotion determined by single subjects rather than by school grades. A very illuminating chapter is the fourth which treats of the difficulties and validity of individual questions in the new-type of tests. All language teachers could learn much about the proper construction of examinations and tests from the study of these pages where among other things the frequency of error and the nature of "good," "bad," and "inverted" questions are discussed.

In the summary and recommendations at the end of this part the author stresses

the following points: (a) the new-type tests being more comprehensive and objective yield more valid, reliable, and comparable measures of achievement; (b) the new-type tests instead of mechanizing teaching serve the best pedagogical interests of students and instructors alike because the efforts of both can be fairly and accurately measured; (c) learning students is prerequisite to teaching them and one of the best means for attaining this end is the use of good examinations; (d) since some students learn more language in one semester than others in four, general policies should provide for these individual differences by the employment of intelligence and language tests in trial periods; (e) the administration of tests and examinations can be justified only on grounds of their constructive usefulness in bringing about homogeneous and well-articulated courses.

PART TWO

In June, 1925, ninety minutes were devoted by second, third, and fourth year language students to a shortened form of the old regents examinations and an equal amount of time to the new-type tests (now known as the Columbia Research Bureau Tests) whose technique is very much like that of the battery administered to the junior high schools in New York City the same year. There were 20,176 cases in French, 8,033 in Spanish, 2276 in German, and 14,081 in Physics scattered throughout the high schools and academies of the state. By halving the old-type examination and applying the Spearman-Brown formula, the reliability coefficients for the different years in French and Spanish ranged from .42 to .79. (This low correlation is due in a measure, we believe, to the shortened form of the regular regents examination.) By using the same process with the new-type test the reliability coefficients for the same subjects and years ranged from .68 to .96; from a careful comparison of the two the author calculates that the new-type test is from one and one-half to two and one-quarter times as reliable as the old.

As criteria for the validity of the new-type tests the school ratings and the review ratings of the regents examinations were put into use and six correlations out of eight were found to be somewhat above .70. The average correlations of the old-type examination with the vocabulary, reading, and grammar parts of the new-type test were .55, .49, and .67, respectively. In regard to the correlations between parts of the new-type tests we shall quote the author's own words:

"The general conclusion is that the intercorrelations of the vocabulary, reading, and grammar tests are high enough to establish the validity of each of them for general measurement purposes, and yet low enough to show that each has a distinctive contribution to make to the measurement of total achievement." (p. 127)

All that has been written above by way of summarizing the results of the experiments in the Junior high schools and the recommendations that follow might be repeated here; instead of doing so, however, we shall merely refer to the conclusion of Part One and call attention to the following outstanding points of the regents experiment:

(a) Taking the new-type test as the criterion, it is shown that the regents overlapping is nearly equal to school overlapping, as for instance about one-third of all the French students in June, 1925, were misplaced by one year or more. This would appear to establish the investigator's claim that there is a fallacy in the assumption concerning the

use of different examinations in the various year classes.

(b) While it is true that the construction of the old-type examinations and the forms of their distribution curves reveal inherent weaknesses, still they should be strengthened, according to Professor Wood, and not abandoned or curtailed. The reviewer believes that the new feature of the regents modern language papers for 1927 and 1928, namely, the introduction of the dictation and oral comprehension parts (which allows a more objective rating than the old scheme of an alternative between a free composition question and oral credit granted by school inspection) ought to improve them.

(c) By means of numerous charts making comparisons in achievement between the results of the old and the new-type tests it is shown that the variable standards of schools are not corrected by regents examinations. This is seen particularly in the case of failing students in some schools having higher marks than those passed in others.

(d) Just as there are inequalities between schools in regard to achievement in language as a whole, so there exists an unevenness in individual schools with respect to vocabulary, reading, and grammar. The causes for this are exposed in a careful analysis of the two types of examinations under discussion, where the author demonstrates among other things that the vocabulary and grammar items of the new-type are arranged in order of difficulty and that the vocabulary sampling is about twice as large as that of the three-hour old-type examination.

(e) Because the new-type test requires only the minimum amount for correction, the cost of the same is estimated to be from one-eighth to one-tenth of that of the old-type.

After affirming that the regents system is sound in principle, that destructive criticisms of it are not justified by the facts, that it is the most advanced educational system of the twentieth century, Professor Wood sums up the results thus:

"The general conclusions from the data of this experiment are that the new-type examinations are roughly twice as reliable and valid as the old-type examinations of equal time allowance; that the new-type examinations afford comparable measures for all classes in a given subject matter and in the same and in different years and thus offers a means of eliminating overlapping of classes and variations in local school standards to a much greater extent than they are eliminated by the old-type Regents examinations.....Our fundamental recommendation, therefore, is that the Regents system should take advantage of the values of the new-type without surrendering any of the *real* values of the old-type examinations..... It seems fair to suggest that the new-type forms of questions be adopted for half of each Regents examination period." (pp. 318-319)

PART THREE

In the last experiment conducted by Professor Wood in June, 1926, the Second Survey of Modern Language Achievement in the Junior High Schools of New York City, 18,870 American Council Beta French Tests, Form B, and 3,940 Spanish Tests, Form B, were administered. Again space permits only of a summary of the most important findings of the author, which follows:

(a) As only slight advantage of the first investigation in respect to reclassification was taken, a careful study of the influence of homogeneity of classes on achievement was impossible. What could be done to throw light on this

problem brought out the fact that there was almost zero correlation between the degree of homogeneity and the amount of achievement, i.e., hardly any relationship at all.

(b) As far as could be learned from the results of the two surveys, a uniform rate of progress in individual students and in classes of the same school was the exception rather than the rule, in fact a matter of chance.

(c) More powerful than any other influence that could be separated was the classroom situation itself.

(d) In spite of factors operating to mask the ideal conditions between homogeneity and achievement and between initial and final achievement of individual students and of classes alike, the second survey manifests the strong tendency toward constancy of achievement rates in individual students. (p. 339)

FREDERIC D. CHEYDLEUR
University of Wisconsin

RESEARCH ABSTRACTS and BIBLIOGRAPHIES

Address all communications relative to research abstracts and bibliographies to Editor A. S. Barr, University of Wisconsin, Madison, Wisconsin.

DOUGLAS, HARL R., AND OTHERS. "Controlled Experimentation in the Study of Methods of College Teaching." University of Oregon Publication, University of Oregon, Eugene, Oregon, 1929.

Five investigations by Douglas and others of the University of Oregon are reported in this bulletin. These investigations are: (1) The value of extra-quiz sections in the teaching of English history. (2) The relative effectiveness of the problem and lecture methods of instruction in principles of economics. (3) Project teaching in a course in educational psychology. (4) An experiment with independent study, and (5) The relative effectiveness of lecture-recitation and supervised-individual methods in the teaching of unified mathematics in college.

GREENE, HARRY A. "Educational Tests and the Classroom Teacher." *University of Iowa Extension Bulletin*, College of Education Series No. 25, University of Iowa, Iowa City, Iowa, 1928.

A bulletin designed to assist the classroom teacher in the use of tests. Materials are presented on the meaning of tests, uses of classroom tests, standardized tests, the selection of tests and their use in the classroom, informal objective tests, their development and use in the classroom, the interpretation of test results, and diagnostic testing and remedial teaching. There is a selected, annotated bibliography.

HOLLOWAY, WILLIAM J. "Participation in Curriculum Making as a Means of Supervision of Rural Schools." *Teachers College, Columbia University, Contributions to Education, No. 301.* Bureau of Publications, Teachers College, Columbia University, New York City, 1928.

A study of curriculum making as a means of supervision. The study was carried on in ten counties. In four of the ten counties the method employed was that of ranking teachers in order of the value of their contributions to courses of study and determining the coefficient of correlation between these ranks and the progress made by their pupils in certain school functions. In the remaining six counties the equivalent group method was used. Teachers of the experimental groups made as much progress in professional spirit, attitude, and teaching skill as did the teacher of the control groups. Children in the experimental groups made, on the whole, slightly greater gains in the functions tested than did children with whom they were compared. Teachers engaged in curriculum making did very much more professional reading than they had ever done before in any one year. Elementary school supervisors were stimulated to greater activity. Communities surveyed by the schools in the experimental group were brought into closer relationship to the work of the schools through the growing power of initiative on the part of the

teacher and the vitalizing effect of the several curriculum making projects.

KOOS, LEONARD V. *Secondary Education in California* (Report of a preliminary survey), California State Department of Education, Sacramento, 1929.

A preliminary report of the survey of secondary education in California. The report discusses the following topics: the growth of secondary education in California, types of districts and schools, the development of higher education as affecting secondary schools, the students, the curriculum, the teaching staff and financial problems.

NATIONAL EDUCATION ASSOCIATION. "Can the Nation Afford to Educate its Children." *Research Bulletin of the National Education Association*, 1928.

The bulletin treats of the following topics: The educational lesson of the war and how we are profiting therefrom; wealth, income, and school support; income tax collections, and school costs in the United States; additional facts about school costs.

PYLE, WILLIAM H., AND MURPHY, HELEN KATHARINE. "Experimental Studies in Pre-School Education." *Bulletin, Detroit Teachers College*, Detroit, Michigan, 1929.

A report of the experiences of Dr. Pyle and Miss Murphy in the pre-school laboratory of Detroit Teachers College.

The bulletin gives the results of two years of experience in three lines of study,—literature, music, and art. In their studies in literature they have tried to find out to what extent children of pre-school age can be trained to get pleasure from books, what aspects or characteristics in stories the children like, and by actually reading many scores of books to them, to discover the merits of the various books and to what ages they are best adapted. Similarly, results are reported for work in music and art. The experiment is a very interesting one and should serve to indicate something of the possibilities of pre-school training.

Report of the Director of the Institute for Child Guidance for the Year Ending June 30, 1928, Institute for Child Guidance, 145 East 57th Street, New York City, 1928.

The report contains discussions of the program, staff, methods of work, etc., for the Institute for Child Guidance.

Report of the Division of Educational Research and Results, Board of Public Education, Philadelphia, Pennsylvania, 1928.

The report contains materials on ability grouping, promotion rates, school progress and elimination, college credit, and teaching efficiency, relation of secondary and higher education, and the results of the educational measurements in the Philadelphia public schools.

Teacher Demand in New York State. (A tentative report of one section of the teacher training study of New York State), State Department of Education, Educational Measurements Bureau, Albany, New York.

The purpose of this report is to determine the yearly demand for teachers in New York State. The report is well documented and as the author indicates 98 percent complete.

WALTER, MILDRED W. "Thrift Education Through School Savings," Cooperative Social Research by Simmons College and the Women's Educational and Industrial Union, Boston, Massachusetts, 1928.

A discussion of school savings systems in Boston and other Eastern cities with

general statistics about the amounts saved in different schools and the distribution of deposits by banks and by systems of administration.

WOODY, CLIFFORD. "Results of the May Testing Programs in the Elementary and High Schools of Michigan." Bulletin No. 119, Bureau of Educational Reference and Research, School of Education, University of Michigan, Ann Arbor, Michigan, 1928.

Data are presented on the Thorndike-McCall reading scale, the Detroit Word Recognition Test, the Stevenson Problem Analysis Test, the Woody-McCall Mixed Fundamentals in Arithmetic, the Wilson Language Error Test, the Iowa Spelling Scale, the Stanford Reading Examination, the Stanford Arithmetic Examination, for the elementary school, and the Kirby Grammar Test, the Van Wagenen Reading Scale, the Wilson Language Error Test, the Columbia Research Bureau French Test, the White Latin Test, the Iowa High School Content Examination, and tests in physics, chemistry, and social sciences for high schools.

EXPERIMENTAL STUDIES REPORTED IN CURRENT ISSUES OF OTHER MAGAZINES

EDUCATIONAL ADMINISTRATION AND SUPERVISION

KLENE, VIVIAN, AND BRANSON, ERNEST P. "Trial Promotion versus Failure." *Educational Research Bulletin*, Los Angeles City Schools, VIII, (January, 1929), 6-11.

A study of the relative merits of trial promotion and repeating. From the data presented the author concludes that the trial promotion group showed greater progress during the succeeding term.

ROBERTS, C. R., AND DAVIS, ROBERT A. "Reading Interests of Teachers." *Educational Administration and Supervision*, XV, (February, 1929), 102-116.

The purpose of this investigation was (1) to determine the amount and character of reading accomplished by teachers; and (2) to analyze some factors which influence their reading. The data for this study were obtained from questionnaires from 240 elementary and secondary school teachers represented in the enrollment in the University of Colorado, Summer Session of 1928. Among the factors studied in this investigation in relation to the character and amount of reading were (1) teaching experience, (2) subjects taught, (3) age, (4) living conditions, (5) professional library, and (6) academic training.

The median number of hours per week of reading per teacher was 14.48. Twenty-nine and eight tenths percent of this time was spent in the field of literature; 35.1 percent in current publications, and 7.1 percent in biography. Ninety-six and seven-tenths percent were engaged, to some extent, in professional reading.

In summarizing the data relative to the factors which influence the character and the amount of reading done by teachers, the author concludes that teaching experience influences the reading only to a small degree, the amount of reading increasing with experience; the subjects taught influence the reading to a marked degree; age influences the reading; and living conditions influenced the amount but not the character of reading. Teachers who had professional libraries read more than those who did not have professional libraries; teachers with only one year of academic training read more

biography and current publications and less professional reading matter while teachers with three years of experience were largely readers of literature.

UNZICKER, S. P. "A Study of Teaching Loads in Junior High Schools in Wisconsin." *School Review*, XXXVII, (February, 1929), 136-141.

A questionnaire study of the teaching load in junior high schools in the state of Wisconsin. Replies were received from twenty-three schools in seventeen cities. A great range in the number of pupils per teacher was found. The median was twenty-two and the mean twenty-four. The number of pupil hours for the middle fifty percent was 593 to 735. The median load was 652 pupil hours a week.

WILLIAMS, LEWIS W. "Supply and Demand as Applied to High-School Teachers." *University of Illinois Bulletin*, University of Illinois, Urbana, Illinois, 1929.

A study of the supply and demand of high-school teachers. The author feels that there is good evidence of an oversupply of high-school teachers in particular lines.

PSYCHOLOGY OF LEARNING

"An Experiment in Learning Vocabulary." *Denver Public Schools Bulletin*, II, (January, 1929), 7-8.

An experimental study of the ability of beginning 1B pupils to learn new words. The results of the investigation briefly stated are as follows:

1. Less than one-half of beginning 1B children learned as many as two of the words presented per day.

2. Absence interfered seriously with the regular work of Grade 1B.

3. The average beginning 1B child knew ten of the fifty words upon entering school.

4. Words vary as to difficulty in learning.

5. Words appearing in groups having like initial letters are not recognized so readily as when they appear in groups of words having different initial letters.

HUTCHINSON, MARK E. "The Correlation Between the Difficulty of Latin Constructions and their Frequency in High-School Latin." *The Classical Journal*, XXIV, (March, 1929), 412-420.

From the data presented it seems that a Latin construction does not automatically become easy for the student because of the fact that it occurs frequently in high-school Latin.

KRUEGER, WILLIAM C. F. "The Effect of Over-Learning on Retention." *Journal of Experimental Psychology*, XII, (February, 1929), 71-78.

This experiment is concerned with the problem of the effect of over-learning upon retention. The author concludes that a certain amount of over-learning, at least fifty percent, is highly economical.

MONROE, WALTER S. "How Pupils Solve Problems in Arithmetic." *University of Illinois Bulletin*, No. 44, Bureau of Educational Research, College of Education, the University of Illinois, Urbana, 1929.

A study of how pupils solve problems in arithmetic. The equivalent group method of experimentation was used. Usable data were secured from 775 sixth-grade pupils, 5902 seventh-grade pupils, and 2579 eighth-grade pupils, or a total of 9,256 pupils from forty-one cities in Illinois. The author concludes that although the data are not entirely consist-

ent, that they appear to substantiate the conclusion that a large percentage of pupils do not reason in attempting to solve arithmetic problems. Instead, many of them appear to perform almost random calculations upon the numbers given. When they do solve a problem correctly the response seems to be determined largely by habit. If the problem is stated in the terminology with which they are familiar and if there are no irrelevant data, their response is likely to be correct.

SYMONDS, PERCIVAL M., AND LEE, BALDWIN. "Studies in the Learning of English Expressions I." *The Teachers College Record*, XXX, (February, 1929), 461-480.

This article is the first of a series of articles on the learning of English expression. It deals with facts of punctuation. Subsequent articles will describe learning for capitalization, vocabulary, constructions, grammatical use, and the subtler matters of style and form. The findings are presented in three ways: (1) a schedule with the total usages, omissions, and errors and comparisons with previous investigations; (2) graphs which show the progress of learning as the quality of composition improves and as pupils move through the school grades; (3) a ranking of omission and error according to frequency for each school grade. A number of tables and graphs are used in the presentation of the data. One of the outstanding facts brought out by this investigation was the wide variation in ability to punctuate in school classes. The middle fifty percent of any school was shown to have a variation equivalent to about five school grades.

EXPERIMENTAL STUDIES OF TEACHING

BURKE, JESSIE D., AND STONE, CLARENCE R. "Relative Effectiveness of Two Different Plans of Training in Silent Reading." *The Elementary School Journal*, XXIX, (February, 1929), 431-436.

The purpose of this experiment is to determine scientifically the relative effectiveness of two plans of training in silent reading, one based on the *Standard Test Lessons in Reading* by William A. McCall and Lelah Mae Crabbs, and the other based upon *Learning to Study Readers* by Ernest Horn and others. The authors conclude that the two methods are of about the same effectiveness for the training purposes studied in this investigation.

ORTON, SAMUEL T. "The Sight Reading Method of Teaching Reading, as a Source of Reading Disability." *Journal of Educational Psychology*, XX, (February, 1929), 135-143.

The author points out how the sight reading method may prove a serious obstacle to certain school children.

REMMERS, H. H., AND STALNAKER, J. M. "An Experiment in Remedial Reading Exercises at the College Level." *School and Society*, XXVIII (December 22, 1928), 797-800.

An intensive study of the gain in speed and comprehension of seven students through the use of motivated remedial speed drills. The results indicate that for approximately three hours of actual reading time spread over a period of nearly two months that there was an increase of approximately 25 percent.

CURRICULUM RESEARCH

HEINLEIN, VIRGINIA K., AND ERVIN, SOPHIA H. "The Voluntary and Independent Use of Reference Books in the Wheeling High School." *The*

School Review, XXXVII, (February, 1929), 142-146.

In this study 841 pupils in the Wheeling High School were asked to answer the following questions: Do you use reference books in the preparation of your lessons if the teachers do not require it? If you do, (a) for what subjects do you use reference books? (b) What books do you use? (c) Where do you find them? (d) How much time do you spend a week in this way? Do you think you understand the use of reference books? Data are presented relative to each of these questions.

The more apparent facts gained from this study are that pupils consciously have a need for reference work, that they use reference books whether or not their teacher instructs them to do so, that they know some of the best and most common reference books, and that they use more reference books and use them to a greater extent as they grow older.

HERRIOTT, M. E. "Objectives of United States History: The Aims of the Elementary School Versus the Expectations of the High School." *School Review*, XXXVII, (February, 1929), 107-116.

An inventory of the objectives aimed at by elementary school teachers and the abilities engendered, and of the expectations held out by high-school teachers. The authors found that there was a gross lack of agreement between elementary-school teachers and high-school teachers of United States history as to the date and name objectives of history. Notwithstanding these discrepancies, the two groups of teachers were in much greater agreement with each other than with the results of one "objective" determination of such objectives. The writer questions the adequacy for curriculum construction of objectives obtained from a consensus of teachers' judgments and from more objective investigations.

EDUCATIONAL STATISTICS

BLISS, ELISHA F. "The Difficulty of an Item." *The Journal of Educational Psychology*, XX (January, 1929), 63-66.

A discussion of a formula for defining the difficulty of an item or single task.

ROWELL, DOROTHY C. "Graphical Comparison of Two Formulas for The Standard Error of a Difference." *Journal of Educational Psychology*, XX (January, 1929), 61-62.

A graphical representation of the statistical results secured from the use of the longer and the shorter formula for the standard error of a difference.

SALISBURY, FRANK S. "A Simplified Method of Computing Multiple Correlation Constants." *The Journal of Educational Psychology*, XX (January, 1929), 44-52.

An extension and simplification of an earlier method presented by writer and Dr. T. R. Kelley in the *Journal of the American Statistical Association*.

WALKER, HELEN M. "Concerning the Standard Error of a Difference." *The Journal of Educational Psychology*, XX (January, 1929), 53-60.

A discussion of the formula for the standard error of a difference showing that

$$\sigma_{x-y} = \sqrt{\sigma^2_x + \sigma^2_y - 2r_{xy}\,\sigma_x\,\sigma_y}$$

and not the shorter formula,

$$\sigma_{x-y} = \sqrt{\sigma_x^2 + \sigma_y^2}$$

should generally be used, the shorter formula being applicable only when $r_{xy} = 0$, which is not usually the case.

NEW TESTS*

An Inventory Test for the Mathematics Needed in High-School Physics. L. R. Kilzer, School of Education, University of Wyoming, and T. J. Kirby, School of Education, University of Iowa.

Plane Geometry Tests, Forms A and B, Maude McMindes, Principal, Senior High School, Hays, Kansas.

English Vocabulary Tests for High-School and College Students. W. T. Markham, Superintendent of Schools, Yates Center, Kansas.

Précis Test. Irvin C. Poley.

A Test in Civic Action. Howard C. Hill and Howard E. Wilson, School of Education, University of Chicago.

Latin Grammar Test. Mark E. Hutchinson, Cornell College, Mount Vernon, Iowa.

Clothing Test. Florence D. Frear, New York State College for Teachers, Albany, New York, and W. W. Coxe, State Department of Education, Albany, New York.

Standard Achievement Test on Principles of Teaching in Secondary Schools. C. W. Odell and M. E. Herriott, University of Illinois.

The Measurement of Socio-Economic Status. Verner M. Sims, Louisiana Polytechnic Institute, Ruston, La.

Seattle Solid Geometry Tests. M. E. Morgan, Roosevelt High School, Seattle, Washington, and W. T. Wait, Department of Psychology, Colorado State Teachers College, Greeley, Colorado, and August Dvorak, School of Education, University of Washington, Seattle, Washington.

Diagnostic Tests in Fundamental Reading Habits. Luella C. Pressey, Department of Psychology, Ohio State University.

*These tests are published by the Public School Publishing Company, Bloomington, Illinois.

MISCELLANEOUS STUDIES

DARINGER, HELEN FERN. "An Objective Measure of Ability to Make Topical Outlines." *Journal of Educational Psychology,* XX, (February, 1929), 112-118.

A description of the construction of a test designed to give an objective measure of the ability of pupils to outline expository paragraphs. The several steps in test construction are carefully presented.

LINDQUIST, E. F., AND FOSTER, R. R. "On the Determination of Reliability in Comparing the Final Mean-Scores of Matched Groups." *The Journal of Educational Psychology,* XX, (February, 1929), 102-106.

The purpose of this article is to point out a statistical fallacy in the method commonly employed to check the statistical significance of gains for matched groups.

WALLIN, J. E. W. AND COLES, CHARLES-ANNA B. "Phonetics Spelling Scales," *Journal of Genetic Psychology,* XXXV (September, 1928), 434-470.

An attempt to devise a predictive test of spelling capacity based upon phonetic words. The author says, "a spelling test consisting of phonetic words would, of course, not constitute such a test, but it would probably constitute a better test of inherent spelling capacity and a better predictive measure of potential spelling ability than would a test composed of non-phonetic words, or a mixture of phonetic and non-phonetic words. Obviously the ability to spell non-phonetic words is more intimately dependent upon familiarity with the words in question than is the ability to spell phonetic words." A percentage scale and a sigma scale for 459 phonetic words is given for grades one to eight.

SYMONDS, PERCIVAL M., AND CHASE, DORIS HARTER. "Practice vs. Motivation." *The Journal of Educational Psychology*, XX (January, 1929), 19-35.

A study of the relative effectiveness of practice and motivation undertaken in the field of learning correct English usage. The experiment was planned so that groups of children were subjected to practice in English usage without motivation and with motivation. Tests were applied before and after the practice and the gains computed. The experiment was carried on in six New York public schools with three sixth-grade rooms from each school. The authors conclude:

"1. The most important single factor in learning as between repetition and the types of motivation used in this experiment is *amount of repetition*. Ten repetitions with only that minimum amount of motivation which habitually operates to cause pupils to do the accustomed work of the schoolroom and to carry out the instructions of the teacher caused more learning than any combination of three repetitions and more powerful types of motivation.

"2. Intrinsic motivation such as manufactured by the devices used in this experiment caused no learning in addition to the practice carried on at the same time. There is strong evidence that motivation from an experienced situation is more effective than motivation from a described situation.

"3. Test motivation caused learning over and above that which could be explained by practice. The value of the test motivation may be estimated as the equivalent of about five sheer repetitions.

"4. There is evidence that there is more gain on tests given later in the series but it is not known whether this is due to the increased effectiveness of the motivation or to better adaptation to the practice."

EDUCATIONAL TESTS AND MEASUREMENTS

BROOM, M. EUSTACE. "The Predictive Value of Three Specified Factors for Success in Practice-Teaching." *Educational Administration and Supervision*, XV (January, 1929), 25-29.

A study of the records of 148 students in the State Teachers College of San Diego. To determine the value of marks in education theory courses, total scores on the Thorndike Intelligence Examination and decide ranks on the trade information section (Part II, Section S) of the Thorndike Intelligence Examination, as means of predicting success in practice-teaching. All of the coefficients of correlation were low, the highest (.296) being between the total scores on the Thorndike Intelligence Examination and practice teaching marks.

KING, THOMAS S. "An Account of the Derivation and Standardization of the Virginia Biology Test." *Virginia Journal of Education*, XXII (December, 1928), 161-166.

The content is based upon the Ruch-Crossman study, validity .68 (28 cases), reliability, .734 to .934 (293 cases). Norms are given.

KUHLMANN, F. "The Kuhlmann-Anderson Intelligence Tests Compared with Seven Others." *The Journal of Applied Psychology*, XII (December, 1928), 545-594.

A detailed comparison of the Kuhlmann-Anderson Intelligence Tests with the Ohio Primary, the Pintner-Cunningham, the Detroit Primary, the National, the Detroit Alpha, the Otis Self-Administering and the Terman group test.

Moss, F. A., Loman, Wm., and Hunt, Thelma. "Impersonal Measurement of Teaching." *The Educational Record*, X (January, 1929), 40-50.

An experiment to measure the efficiency of teachers carried on in twenty-eight Land-Grant Colleges. With the factors of intelligence and previous training held constant, the authors evaluate teachers in terms of changes produced in pupils as measured by tests.

Powers, S. R. "The Correlation between Measures of Mental Ability and Measures of Achievement in Chemistry." *School Science and Mathematics*, December, 1928, 981-986.

The author reports correlations between measures of mental ability and measures of achievement in Chemistry as follows: Powers Chemistry Test and Terman Group I.Q., .31; Cleveland Intelligence Test, .46; Otis Group IB, .67; Otis Scores, .40; Gerry Chemistry, .81; Teacher's Marks, .53; Regent's examination, .35.

Strong, Edward K. "Diagnostic Value of The Vocational Interest Test." *The Educational Record*, X (January, 1929), 159-68.

One hundred and fifty-six seniors of the class of 1927 filled out the Vocational Interest Blank a few weeks before graduation. In January 1928 they reported their occupational choice. Comparison of these two sets of data shows:

Forty-six percent have entered the occupation on which they scored highest in the test.

Twenty percent have entered the occupation on which they scored second highest.

Eleven percent have entered the occupation on which they scored third highest.

Seventy-seven percent have entered the occupations on which they scored first, second or third highest from among fifteen different occupations. Only 18 percent have entered an occupation for which, according to the test, they have no interest.

NEWS and COMMUNICATIONS

Address all research news and communications to Editor Earl Hudelson, College of Education, University of Minnesota, Minneapolis, Minnesota.

The Ohio State University Press announces the establishment of a new educational journal to be known as *The Journal of Higher Education*. It is the hope of the Press that the magazine will be a national medium of publication for materials dealing with the educational and personal aspects of higher education. It will be under the editorial direction of Dr. W. W. Charters of the Ohio State University and a group of widely known associate editors. The business management is under the direction of W. E. Pearce, University Publisher, The Ohio State University, Columbus. The first issue of the *Journal* will appear in September, 1929.

The William E. English collection of original letters, documents, and newspaper clippings on the settlement and development of Indiana Territory, one of the most important collections of source material in early American history, has been given to the University of Chicago by Mr. Willoughby Walling, Chicago banker, grandson of Captain English, for research purposes.

The United States House of Representatives and the Senate have passed as a part of the appropriation bill the following provision:

For all expenses, including personal services in the District of Columbia and elsewhere, purchase and rental of equipment, purchase of supplies, traveling expenses, printing, and all other incidental expenses not included in the foregoing, to enable the Secretary of the Interior, through the Bureau of Education, at a total cost not to exceed $225,000, to make a study of the organization, administration, financing, and work of secondary schools and of their articulation with elementary and higher education.

Provided, That specialists and experts for temporary service in this investigation may be employed at rates to be fixed by the Secretary of the Interior to correspond to those established by the Classification Act of 1923, as amended, and without reference to the Civil Service Act of January 16, 1883.

The plan of expenditure contemplated in this action is the use of $50,000 during the year 1929-30, $100,000 during the year 1930-31, and $75,000 during the year 1931-32.

Superintendent William J. Bogan of Chicago has organized an Advisory Council of eighty representative citizens for the purpose of bringing the best thought and experience of the city to bear upon the numerous problems of educational policy. At the second meeting of the Council, on January 24, Superintendent Bogan addressed the members in part as follows:

"I am sure you all understand now, if not before, that this organization is made up of a truly representative group of people interested in the education of our children. Long ago we found that merely to hire a superintendent and a group of associates, principals, and

teachers, and tell them to go to work in the education of the children of Chicago would not turn the trick. Every citizen of this great city ought to be interested in the education of the children and, therefore, I organized this group some months ago in order that all the knowledge possessed by specialized organizations interested directly or remotely in the education of children might be made to focus upon the office of the Superintendent of Schools so that he might transmit that information, that experience, to all those who have the task of doing the work directly."

Assistant Superintendent James E. McDade announced that recommendations of the highest value were being received from the committees on Character-Building Instruction, Child Guidance, and Health Preservation, and that other committees are being organized as occasion demands.

More women specialize in English than in any other field of study and more men in Economics, according to a study just made at Radcliffe College of the choices of Radcliffe students and men of the same classes at Harvard. English stands second at Harvard and Economics thirteenth at Radcliffe. The comparison has especial interest because the curriculum is, with a few minor exceptions due to the difference in interest between the sexes, the same at both institutions and because all the Radcliffe teaching is done by members of the Harvard Faculty of Arts and Sciences.

Twenty-eight percent of the women at Radcliffe are "concentrating" in English, as against 16 percent of Harvard men. Seventeen percent of the students at Harvard, on the other hand, have chosen to "concentrate" in Economics, while less than three percent of Radcliffe students prefer that business-like subject.

The first ten subjects on the Radcliffe list, which make up 80 percent of the choices, are, in order of preference, English, Romance-Languages, Fine Arts, History, History-and-Literature, Mathematics, Biology, Psychology, Classics, and Government. At Harvard the first ten are Economics, English, History, Romance-Languages, Chemistry, Fine Arts, Government, History-and-Literature, Bio-Chemical Sciences, and Mathematics.

There are three subjects among the first Radcliffe ten which do not appear on the Harvard list: Biology, Psychology, and Classics. There are also three among the first Harvard ten which fall outside the first ten of the Radcliffe choices: Economics, Chemistry, and Bio-Chemical Sciences.

Dividing the fields of study into cultural subjects on the one hand and scientific and practical subjects on the other, an interesting difference is discovered. At Harvard the choices in the cultural and the scientific and practical groups are almost exactly even—approximately 50 percent in each. At Radcliffe, however, nearly three times as many preferences are expressed for the cultural subjects—72 percent as against 28 percent for the scientific and practical fields.

It further appears from the study just made that more than half of the Radcliffe women have elected to be candidates for honors, while a few more than one-third of the men at Harvard have chosen to study for honors. The exact figures are 57.2 percent at Radcliffe and 35 percent at Harvard.

Professor L. L. Thurstone of the Psychology Department of the University of Chicago and Mr. Richard W. Jenkins, M.D., of the Chicago Institute for Juvenile Research have completed a two-year analysis of the intelligence and health

record of the children of fifteen hundred Chicago families. Their results show that "older children in each family tend to be less intelligent, on the average, and more susceptible to disease and insanity than the younger children." Other findings are that "the age of the parent at the time of the child's birth does not affect the child's intelligence; that discrepancy between the ages of the parents likewise has no effect; and that the level of intelligence of the children of small families is higher than that of large families."

Miss Minnie Steckel, graduate student, is making a similar survey among ten thousand school children in Sioux City, Iowa, and preliminary reports indicate that she is getting similar results.

Educational Measurement as a College Course

"Educational Measurement", the name applied to a course offered in our colleges and universities is usually a misnomer. At the outset, it must be borne in mind that we are dealing with the first course (an introductory one) in the study of measurement, a course given to seniors in college or in the first year of graduate study in education.

The over-evaluation of statistics on the part of the teachers of administration and secondary schools has led not only to the insistence on their part that the course be related to their field, but also has brought about unnecessary quibbling concerning formulas and methods of interpretation wherein the professor of educational measurements chanced to disagree with them. Furthermore it has led to that over-emphasis on statistical treatment of scores, etc., that is prevalent in many institutions.

Either the above viewpoint has been taken by many administrations or, worse, a negative, depreciatory one has been held concerning measurements courses in general. This attitude is exemplified by the remark of a secondary education professor to the effect that "All the statistics needed (hence all that his pupils will gain from educational measurements) by the educator can be taught in one week-end of not too intensive study." Such attitudes make the work of the educational measurement professor none too easy.

A second attitude which is even more prevalent than the above is that a course in educational measurement should teach mainly the testing of intelligence and the use of intelligence tests, an attitude held by many psychology teachers and clinical workers. This practice is very common at the present time in many of our institutions and is probably responsible for the horde of half-informed mental testers and psychometrists that infest our school systems. Extension classes, particularly, are guilty of such practices.

Intelligence testing is an art in itself, a technique difficult to attain even in a systematic course dealing with one specific test, such as the Stanford Revision of the Binet Tests; and the thoroughness demanded, in our estimation, calls for a separate course. In liberal arts colleges we have seen the ill-fated attempt on the part of instructors to teach accomplishment tests, group intelligence tests, and the Binet Test in a course already overloaded with statistics, graphical representation and interpretation of data, etc. The smattering of information gained by the student was of little value to him as a teacher, nor did it increase his respect for professional courses of any kind.

Again, the educational measurement course should not place its main emphasis on accomplishment and character tests with the immediate application of

such tests to a school system. This, then, is a third over-emphasis that many times is resorted to in this course.

A fourth phase of this subject that is sometimes taught with little or no regard for allied aims and functions is the statistical treatment and interpretation of test scores. This is often the result of the over-pressure exerted by other departments as intimated above; but more often it is due to the fact that the instructor himself, working on research problems and frequently contributing to psychological and educational magazines, entailing as this does continual use of statistical measures of representation of data, acquires an inflated impression of the value of statistics *per se*. Such an instructor is particularly prone to overlook the makeup of his class and the probable sphere that these people will occupy in the educational world in the future, and hence over-estimate the functional value of statistics for these men and women.

Above all, a study of the derivation of complicated formulas should be avoided in a course of this nature, and should be left to an advanced course for the few men who will have need of such knowledge.

We have outlined four of the most prevalent and most hampering causes of over-emphasis that may be found in the various courses of educational measurement. What, then, should comprise the instruction in such a course? An analysis of the student group in such a course may help to answer the question. In the average course in educational measurement we generally find the following types of students:

1. Those who are to be educational administrators, heads of public secondary schools, etc.

2. Future teachers (male and female) of educational measurement courses in colleges and normal schools.

3. Private school teachers and executives.

4. Psychologists, interested in the testing phase merely for its informative value; i. e., as a distinct aspect of their subject.

5. Students whose future occupation will be one of the following:

   a. School psychologist.
   b. Social service worker.
   c. Juvenile court assistants.
   d. Institution workers in schools for mental defectives, special disabilities schools, institutions for the deaf, blind, crippled, etc.
   e. Mental hygienists.

6. Graduate students who are interested primarily in the statistical treatment and representation of educational data.

7. Those who will teach other subjects such as English, Latin, etc., in secondary schools and colleges, and who are interested in the application of testing technique to their own fields.

Not only do we find a great diversity in the aspiration and aims of the students but we also find gross differences in their academic equipment. We have graduates of Normal colleges, technical schools, liberal arts colleges, sectarian colleges, etc. Some know little or no psychology, many have but a high-school training in mathematics, and the professional training in Education differs greatly in each case. Again, the difference in age and in teaching experience makes for a heterogeneous working group.

Faced with such a group as we have outlined above—and we believe we have not overdrawn the divergence existing

within the average class—the instructor obviously cannot meet the needs and fancies of all. To do this would require more than one course and more than one semester's time. This state of affairs, however, has led many instructors in this field to concoct a veritable potage as an academic offering.

There seems to be but one way out and that is to make the course a survey course, introductory to the more specific fields of testing that will interest the educational psychologist only. For *general information* concerning testing is needed by all the people mentioned above, nor would we exclude them; but it does necessitate a change in the content of such a course and in the procedure adopted by the instructor himself.

Suggested Study on the Relation of Training of Teachers in Content Subject Matter to the Effectiveness of Instruction

I. Reasons for the Proposed Study

In a recent study[1], Professor Buckingham shows that for Ohio there existed in 1923-1924 practically no correlation between the training of secondary school teachers in content subject matter and the subjects taught by such teachers in the high school. Correlating the frequency with which each subject was studied with the frequency with which it was taught, he found the following correlations:

Major subjects with teaching subjects:
$r = +.13$

Minor subjects with teaching subjects:
$r = -.07$

Major and minor subjects combined with teaching subjects:
$r = +.14$

These low correlations are cause for concern. At the present time it is at least desirable that a student who is preparing to teach a subject in high school should have considerable training in college in the content field. There are those who maintain, and so far no one has disproved their contentions, that the essential thing in training a teacher is to have him become a master of the so-called higher content phases of that subject. To teach mathematics in high school, have the student-in-preparation major in mathematics,—that is, study college algebra, analytical geometry, trigonometry, the differential calculus, the integral calculus, theory of equations, —and the knowledge thus gained will be the best insurance that he can then teach elementary algebra and plane demonstrational geometry with effectiveness to high-school pupils.

It is possible and feasible to test this contention, and, with the science of education developing techniques with alacrity, it will not be long until the functional effects of teacher training will receive more than an argumentative evaluation.

To pursue the mathematics illustration further, if majoring in college mathematics does function in the direction of getting better pupil achievement in high-school mathematics, or if it functions toward getting the desired degree of achievement in less time, then, for the sake of increased mathematical ability in the population at large or for the sake of economy of time in developing mathematical abilities desired in a democracy, the requirement by a college of a college major in mathematics for those who are preparing to teach mathematics in high school becomes essential.

If, on the other hand, it is found that some of the college mathematics courses

[1] Buckingham, B. R. *"Supply and Demand in Teacher Training." Bureau of Educational Research Monograph*, (Columbus: Ohio State University, Vol. II, March 15, 1926), p. 108.

do function and that others do not, then, for training purposes, more emphasis may be placed on those courses that do function and less on those that do not. If it should happen that the calculus or the theory of equations function not at all or in small degree in securing better pupil achievement in high-school algebra and high-school geometry, then the calculus and the theory of equations are non-essentials to the curriculum for high-school mathematics teachers.

Again, if it be found that certain parts of college algebra function to a greater degree in securing better pupil achievement in high-school algebra than do certain other parts, then it may be desirable to place less emphasis for purposes of teacher training on certain parts of college algebra and more emphasis on certain other parts. It, of course, is obvious that there are certain connective materials which might not show a direct relation to pupil achievement but which may, nevertheless, be basic or be prerequisite to topics of high value in teacher training. These materials would need to be included in the teacher-training course in the natural order of sequence.

If the functional effects of special training in securing better teaching would be tested, the following practical outcomes might be possible:

1. In providing a well-balanced curriculum for a high-school teacher, those in charge of teacher training institutions could effect a more intelligent apportionment of a student's time between the content field of his specialty and other fields which may occupy essential places in the curriculum.
2. The emphasis placed upon the various divisions and topics within the field of the special content could be varied as the conditions would seem to warrant.
3. It would reveal the importance of regulations for the training of teachers for specified positions and emphasize the necessity of establishing agencies for the placement of properly trained teachers in specified positions.

II. The Procedure Proposed

The following procedure is proposed as a method which gives promise of answering in some degree the question of the extent to which the training in mathematics received in college by high-school teachers of mathematics functions in their teaching.

1. Select 200 persons who are teaching mathematics in high schools of over six hundred enrollment, to coöperate. These teachers should be selected to fill the following conditions (minimum).
   - A. 40 with M.A. degrees in mathematics from six universities giving similar training in mathematics.
   - B. 40 with A.B. degrees with mathematics as a major.
   - C. 40 with A.B. degrees with mathematics as a minor.
   - D. 40 with A.B. degrees with one year of mathematics.
   - E. 40 with A.B. degrees with one term or less in mathematics.

2. Procure data on the two hundred teachers as to
   - A. Training
   - B. Experience
   - C. Rating on college entrance ability test
   - D. Scholastic records.

The two hundred teachers who make up the five groups A, B, C, D, and E, under (1) above may need to be selected from a larger number. The final study may be made from two hundred teachers

selected from the larger number so that the factor of training, except in mathematics, will if possible be approximately equal in all five groups. It can be made equal at least for the last four groups. Years of experience and other factors will also be equated. In other words each group of teachers will be selected so that all factors save one, training in mathematics, will be as nearly constant to all five groups as possible. Each group of teachers will be homogeneous with respect to all measurable factors but will differ from every other group with respect to the amount of training received in college mathematics. Having data on the teacher as outlined above we must next contrive some way to measure the functional effects of the teaching. We proceed, therefore, to step (3).

3. Construct for the special purposes of this study, achievement tests in high-school algebra and geometry. These will be designed so as to test the achievement of the pupils in a manner such that the results, when analyzed, will afford two types of comparisons.

First, the results will afford a comparison of the gross achievement of pupils who are taught by teachers who have had different amounts of training in college mathematics, but who are in other respects homogeneous.

Second, the results will afford a comparison of the functional effect that certain types of itemized content may have upon the achievement of the pupils classified as were their teachers under the various groups.

The actual construction of the tests will need to be based upon a careful content analysis. Only such content as contains determinable elements will lend itself to functional analysis. Mathematics meets this condition. Algebra taken as illustration may help to make this point clear.

High-school algebra may be analyzed into its constituent elements. These elements may be classified under appropriate headings, such as definitions, notations, equations, factoring, and the like. Under definitions would then be listed the various types. Subject matter thus specifically listed would be considered as elements. Thus the elements with the frequency of recurrence would be noted.

College courses are likewise analyzed, preserving in such analyses, as far as possible, the classifications used when analyzing the high-school course. The technique is illustrated in a number of studies that have been made on the amount of overlapping in the content of a field in various educational units.

The elements in the high-school course will reappear in the college courses. In some cases these will appear in practically the same context. Such materials may be classed as review. The same elements may, however, appear in new relationships or may recur with variant concomitants. The final analysis may reveal the elements of high-school algebra arranged in the order of their frequency of occurrence. Similar treatment may be made of college algebra. We may distinguish in this analysis between elements in high-school algebra that recur in identical form in higher mathematics courses and those that recur with variant concomitants. Thus both the form and the amount of recurrence can be recorded and the place of recurrence noted.

The tests must be constructed in light of what the analysis reveals. The test, or it may be a battery of tests, will need to encompass a number of situations. Thus it may include a group of items which recurs in all of the courses; another group which scarcely recurs at all; or another group that tests for materials found in one college course but

not in any other. In other words, an exacting content analysis which reveals the extent and nature of recurring elements will serve as a guide to the test construction. Once the tests are constructed the fourth and last step will be relatively simple.

4. With the aid of these tests, procure data upon the achievement of all pupils selected for purposes of the study. Using a technique which may well be somewhat similar to one reported by the writer[2] in another study, one may make the following studies and a number of others as well.

A. Compare the achievement of pupils in high-school algebra, plane geometry, advanced algebra, and solid geometry, who have taken these subjects under teachers with varying amounts of training in college mathematics. The personality factor which is always a difficult factor to control in this type of study may partly be taken care of through the treatment of the data on intelligence and scholastic records. Personnel blanks in the universities may also help some.

B. Compare the achievements of pupils on elements of materials found to recur with greater or lesser frequency in courses included in the teacher training curriculum, with the achievement of pupils on the same elements but under teachers whose training curriculum did not offer recurrence in these elements. Similarly, make comparisons of the achievement of pupils who were under teachers whose training curriculum afforded different types and frequencies of recurrence.

Because of brevity some of the steps outlined above may not seem clear. Most of the factors in a study like this can be made nearly constant and large group comparisons of pupils under teachers differently trained seemingly should afford interesting and somewhat significant results. The technique once it is used in one subject field will reveal a way to make similar studies in other subject fields. Teacher training programs may therefore for the first time be partially evaluated.

III. Suggested Limitations

The investigator will meet with the same difficulties ordinarily found in studies which include the human elements of personality of teachers, skill in presentation, nature of selection, and the like. Records mentioned above would help to isolate these factors.

The effect of so-called general training could not be measured. Training in professional education could partly be controlled, but possibly the factor of general education would have to be assumed as constant within all groups.

In studying pupil achievements, concomitant factors such as ideals, attitudes, and the like would need to be disregarded.

In spite of its limitations the study gives considerable promise of affording interesting and, not unlikely, surprising outcomes. We are most meticulous in recommending teacher-training programs and are considerably concerned when superintendents place teachers who are not specially trained in positions where specially trained teachers are readily available. In reality, we may have little

[2] Hughes, J. M. "A Study of Intelligence and of the Training of Teachers as Factors Conditioning the Achievement of Pupils," *The School Review*, xxxiii (March and April, 1925), 191-209; 300-302.

ground for contending that the training of the one will function more effectively as shown in the form of better results than will the supposedly inferior training of the other.

We should find out. Possibly some one properly qualified and with plenty of time and some money will feel encouraged to attempt it.

J. M. HUGHES
Northwestern University

THE USE AND MISUSE OF FIGURES OF SPEECH IN EDUCATIONAL WRITING

In a recent educational journal of national circulation there occurs an article entitled, "The Vital Organs of the Mind."[1]

The following sentences are taken from the first few paragraphs of the article:

> "In the running of the mind, one must perforce be not only his own chauffeur, but also, in large measure, his own repairman."
>
> "Each of us should be able to recognize the cylinders of his mental motor, control them and keep them from racing, and clean an occasional dirty spark plug."
>
> "The heart of the human personality, its 'cylinders,' are the primary emotions, such as fear, anger, elation."

Surely man is fearfully and wonderfully made! The title of the article leads one to look for an analogy in which the "mind" is to be compared with a living organism, but it turns out that the "mind" is an automobile with a chauffeur, a repairman, a motor, racing cylinders, and even dirty spark plugs. Then to return to the first figure, the human personality seems to have a heart. Perhaps the heart constitutes the "vital organs" of the title; or are we to have the rest of the vital organs later—the lungs, the brain, the stomach, the liver, and maybe a gall-bladder?

When I read serious educational articles like this, written in such highly figurative language, my mental stomach seems a bit unsettled—or perhaps it is my spleen—or should I say my ignition system? As mental chauffeur I have difficulty holding my car in the road, and I am afraid that as a mental repairman I shall have to spend an undue amount of time under the car. So I am running in for repairs—that is, to be advised that such figures are good pedagogical devices for readers too ill-informed to be told the truth in other terms; or else I am seeking to lend support to a movement for having the road repaired—that is, to see if something cannot be done to discourage the use of figures in this way.

Seriously—laying aside stomach, heart, cylinders, and chauffeurs—this note is written to raise the question whether men of scientific training should resort to such highly figurative descriptions of mental mechanisms and mental processes, even when dealing with the untrained. Do they not carry seriously misleading implications? Cannot the purpose of the writer be accomplished by descriptions that are more rigidly scientific and yet within the ability of the probable reader? Is it not possible to carry the main point and at the same time stimulate correct "marginal responses?"

It is clear, of course, that the criticism of this note is largely destructive. One has a right to demand that the writer of the criticism give an illustrative presentation that is correct, understandable, and appealing. That task, however, I shall leave to someone more skillful. Can it be done?

H. T. MANUEL
University of Texas

[1] The exact reference is omitted, for the criticism is in no sense a personal one.

The Sangren-Woody Silent Reading Test

Description of the test

The Sangren-Woody Reading Test, recently prepared by Dr. Paul V. Sangren, Director of Educational Measurement and Research, Western State Teachers College, Kalamazoo, Michigan and Dr. Clifford Woody, Director of Educational Reference and Research, University of Michigan, is so constructed as to measure seven rather distinct phases of silent reading ability. The seven phases of silent reading ability measured by the various parts represent the more important ones so far as may be determined by the generally accepted aims of silent reading instruction. A description of the various test parts follows:

1. Word Meaning—the extent of the child's silent reading vocabulary.
2. Rate—the child's rate of reading a continuous selection.
3. Fact Material—the efficiency with which the child can read and get facts from the material.
4. Following Directions—the efficiency with which the child can read and follow directions on the basis of the material read.
5. Total Meaning—the efficiency with which the child can read and grasp the total meaning of the paragraph read.
6. Central Thought—the efficiency with which the child can read and detect the central thought of the paragraph read.
7. Organization—the efficiency with which the child can read and organize logically the ideas contained in the series of paragraphs read.

Assumed criterion of achievement in silent reading

Assuming as a criterion of silent reading ability a composite score derived from the following measures the correlations included in Table I were obtained.

Stanford Achievement Test: Paragraph Meaning
Stanford Achievement Test: Sentence Meaning
Stanford Achievement Test: Word Meaning
Thorndike-McCall Reading Scale
Burgess Silent Reading Test
Monroe Reading Test (Revised)—Rate
Monroe Reading Test (Revised)—Comprehension
Teachers' Reading Marks (Averaged for 3 months)

TABLE I

Correlations Between Criterion Scores and Scores on the Separate Parts of the Sangren-Woody Reading Test*

	r	P E
Criterion and Word Meaning	.888	.008
Criterion and Rate	.754	.017
Criterion and Fact Material	.901	.007
Criterion and Following Directions	.710	.020
Criterion and Total Meaning	.753	.017
Criterion and Central Thought	.688	.021
Criterion and Organization	.610	.024
Average coefficient	.76	
Criterion and Total Score	.897	.007

*Results based on 310 cases, including 66 4B-, 59 5B-, 62 6B-, 53 7B-, and 70 8B pupils.

Reliability of the test

Table II gives coefficients of reliability obtained from the application of two forms of the test to 100 VIB and 110 VIIB pupils.

TABLE II

Reliability Coefficients for the Parts of the Sangren-Woody Reading Test

Name of Part	Group Included VIB		VIIB	
	r	P E	r	P E
Word Meaning	.866	.017	.906	.012
Rate	.900	.013	.903	.012
Fact Material	.875	.016	.777	.027

Follow Dir'ns	.828	.021	.809	.023
Total Meaning	.877	.015	.870	.016
Central Thought	.894	.015	.868	.016
Organization	.881	.015	.919	.010
Average conf.	.88		.86	
COMPOSITE	.937	.008	.945	.003

UTILITY OF THE TEST

The Sangren-Woody Reading Test may well be put to the following uses:

1. To determine the weaknesses and strengths of individuals and groups in the various aspects of silent reading. This may be indicated by means of the profile chart.
2. To determine more adequately the general silent reading ability of individuals and groups through the use of the total score in comparison with the norms of achievement.
3. To serve as a basis for instructional procedure in building up practice exercises in the various important aspects of silent reading.
4. To serve as a basis for sectioning children into more homogeneous groups for purposes of reading instruction.

PAUL V. SANGREN

Western State Teachers College, Kalamazoo, Michigan

AGE OF MATRICULANTS FROM JUNIOR COLLEGES

The following three distinct types of public junior colleges are found in California: (1) independent junior colleges, organized in separate junior college districts, with board of education, faculty, administration, and physical equipment usually independent of the high school; (2) junior college departments of high schools, in which two years of work are given in direct connection with the high school, with the same board and administration, and with faculty and plant more or less in common; (3) junior college departments connected with six state teachers colleges, where two years of college work is given under the regular administration and faculty of the institution.

In the past five years Stanford University has received by transfer students from nine independent junior colleges, from eight of the high-school type, and from six of the teachers college type, all in California; and in addition from fourteen non-California junior colleges, both public and private, located in Alabama, Arizona, Illinois, Massachusetts, Missouri, New Mexico, Texas, Utah, and Washington. The total number of such matriculants has been 510. Of this group there were 317 who had completed two full years of junior college work before transferring, and who accordingly were given upper division or junior-class standing when they matriculated at the University.

Significant data regarding age at entrance of this group of 317 students is summarized in Table I.

TABLE I

AGE AT MATRICULATION OF JUNIOR COLLEGE ENTRANTS AT STANFORD UNIVERSITY, 1923-24 to 1927-28

Group	Number	Mean Age (Years)	P.E. of Mean
Total	317	20.52	0.06
By Type of Junior College			
Independent District	141	20.50	0.09
High-School Type	25	20.20	0.21
Teachers College Type	121	20.71	0.09
Non-California	30	20.08	0.19
By Year of Entrance			
1923-24	30	20.93	0.19
1924-25	56	20.43	0.14
1925-26	62	20.71	0.13
1926-27	85	20.37	0.11
1927-28 (autumn only)	84	20.43	0.11
By Sex			
Men	264	20.60	0.06
Women	53	20.09	0.14
Stanford Groups for Comparison			
All Freshmen (Maximum age 29)	679	18.55	0.04

All Advanced Standing (Maximum age-51)	409	22.20	0.17
Advanced Standing (Maximum age-29)	385	21.17	0.13
All Graduate Students (Maximum age-60)	402	31.04	0.10

The distribution of the ages of the entire group of 317 junior college entrants is shown in Table II.

There is slight difference in age between entrants from the different types of junior colleges. The only significant difference is between the teachers-college type and the non-California group, the latter being two-thirds of a year younger. (Difference, 0.63; P. E. of difference, 0.21).

There seems to be a slight tendency for the age to decrease in the more recent years. The maximum difference is over a half year between the 1923-24 entrants and those for 1926-27 (Difference, 0.56; P. E. of difference, 0.22). There is a difference of a little more than a quarter of a year between the average of the first three years and the last two. (Difference, 0.29; P. E. of difference, 0.12).

TABLE II

AGE AT MATRICULATION OF 317 JUNIOR COLLEGE ENTRANTS AT STANFORD UNIVERSITY 1923-24 TO 1927-28

Age at last Birthday	Number
16	1
17	8
18	25
19	73
20	110
21	54
22	26
23	13
24	4
25	2
26	0
27	0
28	0
29	1
Mean age	20.52 years
Median age	20.45 years

Of 317 junior-college entrants, 1 was over 25 years of age.
Of 679 freshmen entrants, 4 were over 25 years of age.
Of 409 advanced standing entrants, 51 were over 25 years of age.

The women are significantly younger than the men. (Difference, 0.51; P. E. of difference, 0.15).

The entire group of junior-college entrants (with two years advanced standing) are almost exactly two years older than entering Stanford freshmen. Unless elimination affects age, the freshmen when they become juniors will be of the same age as the junior-college group. (Difference, 0.30; P. E. of difference, 0.07).

The entire group of junior-college entrants are one and two-thirds years younger than the group of advanced standing students (Difference, 1.68; P. E. of difference, 0.15).

The 24 advanced standing students whose ages exceed 29 years have an undue effect on the mean. If they be eliminated, the junior-college group is still two-thirds of a year younger than this more nearly comparable group of advanced standing students. (Difference, 0.65; P. E. of difference 0.14).

The junior-college group has a fairly symmetrical distribution of ages. The median and mean differ but slightly.

The facts reported above point to a somewhat different conclusion from that reached by Koos based upon a comparison of ages of freshmen in sixteen public junior-colleges with ages of freshmen from a distance at the University of Minnesota. He finds a difference of eight months and concludes "that the public junior-college tends to prevent delay in entering upon college work."[1]

Over ninety percent of the freshmen at Stanford come "from a distance" since it is not located in a metropolitan area like the University of Minnesota. Thus

[1] Koos, L. V. *The Junior-College Movement.* Boston: Ginn and Co., 1925, p. 130.

the average age of all freshmen and those from a distance would differ but slightly, if at all. The data summarized above show that students entering Stanford as freshmen averaged 18.55 years of age, while those entering as juniors after graduation from the junior colleges averaged 20.52 years of age. Two years earlier, when they were probably entering junior-college as freshmen, their average was 18.52 years. It is true that these figures do not apply to all junior-college freshmen, but only to those entering Stanford University after graduation from the junior-college. To this extent the conditions are not comparable with the data reported by Koos. But for the groups considered it is clear that attendance or non-attendance at a junior-college is not a significant element of difference as regards age.

WALTER C. EELLS
Stanford University